Signal Processing for the NII

Workshop/Panel Report

Ballston, Virginia
18–19 August 1994

Edited by
Robert M. Gray

Sponsored by
The National Science Foundation
Circuits and Signal Processing Program
Microelectronics Information Processing Systems Division
Computer and Information Science and Engineering Directorate

Telephone Device for the Deaf

The National Science Foundation has telephonic device for the deaf (TDD) capability which enables individuals with hearing impairment to communicate with the Division of Personnel and Management for information relating to NSF programs, employment, or general information. This number is (703) 306-0090.

Electronic dissemination

You can get information fast through STIS (Science and Technology Information Systems), NSF's online publishing systems, described in the "STIS Flyer" at the end of this document. This document is also available via anonymous ftp at isl.stanford.edu in the directory /pub/gray/reports.

Ordering by Electronic Dissemination

If you have access to Internet, you may order publications electronically. Internet users should send requests to pubs@NSF.gov. In your request, include the NSF publication number and title, number of copies, and your name and complete mailing address. Printed publications also may be ordered by FAX (703-644-4278). Publications should be received within 3 weeks after receipt of request.

Contents

Chapter 1

Executive Summary

1.1 The Workshop

The field of signal processing (SP) encompasses the theory and practice of algorithms and hardware that convert signals produced by natural or artificial means into a form useful for a specific purpose. The signals might be speech, audio, images, video, sensor data, telemetry, electro-cardiagrams, or seismic data; the purpose might be transmission, storage, display, diagnosis, classification, segmentation, or interpretation. SP is one of the basic enabling technologies for the evolving National Information Infrastructure (NII) — SP has played a fundamental role in NII's development and use, and is continually changing in focus to meet many new challenges presented by the NII.

In August 1994 a two-day workshop/panel was held at the National Science Foundation to assess the current state of SP in the NII, to identify important issues and trends of research and development of SP theory and applications for the NII, and to make recommendations for research and development. The discussions focused both on the development of enabling technologies — the underlying tools that are useful for a wide variety of applications — and on several specific applications with strong SP components.

This report summarizes the proceedings of that meeting and of Email conversations before and after the meeting. The participants included active researchers from academia, industry, and government in a variety of fields, including electrical engineering, computer science, geographical information systems, planetary science, medicine, and education, reflecting the variety of current and potential users of the NII.

1.2 Overview

SP has been a key enabling technology for the NII, contributing to efficient and reliable acquisition, storage, transmission, and reconstruction of speech and audio, images and video, sensor and measurement data, and other information bearing signals, and to efficient use of various delivery channels including phone lines, coaxial cable, optical fiber, and wireless. Key standards have been a direct result of on-going research in both academia and industry. The ubiquitous nature of SP makes it transparent to most users, yet the NII could not function without efficient methods for converting, compressing, coding, equalizing, detecting, encrypting, communicating, storing, searching, interpreting, and displaying image, audio, and other data. Familiar applications include medical image processing, geographical information systems, consumer multimedia, remote sensing and planetary research, and education. The field has reached the point where the technology for generating, acquiring, storing and transmitting data in all these applications is advanced, stable, and increasingly useful. Further research is likely to lead to incremental but worthwhile improvements in traditional methods, but the emphasis of SP research and development is shifting from the communication and storage of signals without reference to their content to problems of integrating, digesting, and comprehending information. These new algorithms will be strongly application dependent and their development into practical products will require active cooperation among the developers of the SP algorithms and the specific user community. The collaboration required among diverse groups of widely varying training and technical sophistication presents an enormous challenge and opportunity to the NII development and user community.

1.3 Findings and Recommendations

The findings and recommendations highlighted below were developed over two days of active discussion at the workshop and through extensive Email exchanges in the weeks surrounding the event. Chapter 2 expands these findings, while Chapters 3 and 4 provide their context. The motivation and justification for the specific findings and recommendations are treated in some depth in Chapter 5 and exemplified in Chapter 6.

Grand Challenges

One general goal and two general areas provide common themes underlying many of the specific recommendations.

Universal Access for Education Provide useful information and applications to all users, including those with minimal facilities and training.

Tools to Support Content Analysis Promote aggressive interdisciplinary research in methods that extract content from signals by performing modeling, identification, classification, segmentation, recognition, and other forms of data analysis. These challenges can be met only by researchers capable of spanning a broad intellectual and scientific range that includes SP, applied mathematics, computer science, statistics, linguistics, and systems/control theory.

Data Quality Validation Develop acceptable and trustworthy protocols for validating the utility and quality of signals altered by SP and develop means of predicting utility and quality from easily made quantitative measurements. This can involve simulations of user practice such as simulations of radiologists making clinical diagnoses, statistical analyses of differences in decisions or measurements made on processed and unprocessed images, human perceptual models, subjective testing, and the development of improved quality metrics for signals that are seen and heard.

Success in responding to these challenges will require the active participation of the intended user communities (e.g., educators, scientists, and radiologists) in developing truly useful SP tools.

Cooperative Research in Key Areas

Finding:

Theory and algorithms for SP must be developed in the context of intended applications and in close cooperation with users if the algorithms are to reflect the needs of the user community. This involves interdisciplinary research, often involving many organizations and institutions as well as researchers of widely varying backgrounds. Specific key research areas are identified by the panel are detailed in Section 2.3.

Recommendation:

Genuine, interdisciplinary and cooperative efforts for research on SP for the NII among academia, government and industry should be initiated and supported to the greatest extent possible. Our suggestions include initiatives for developing asymmetric multiresolution image processing tools for medical, educational, or scientific information systems.

Standard Databases

Finding:

Theoretical development of SP algorithms can often take place using easily available data for preliminary tests, but genuine comparisons among competing algorithms and validation of quality and utility require universally available standard data sets. For example, computer aided diagnosis of CT scans or feature extraction from telemetry data may behave quite differently on some data sets than on others. Hence common standard databases enhance research and permit comparison between competing schemes.

Recommendation:

Support the establishment of standard databases for SP research. This can be done either through financial support to individual institutions (as may be necessary for expensive databases such as high resolution digital mammograms) or by providing a clearing house of publicly available databases, reachable via the NII. The data types considered in Chapter 6 provide specific examples.

Standards for Tool Validation

Finding:

Not all new SP algorithms developed over time can or should be made readily available for NII use. *De facto* standards exist simply because they work well and no clearly superior algorithms performing similar or better functions have appeared. Nonetheless, it will be useful to develop standards to validate new algorithms, that is, to demonstrate *quantitatively* that they perform as claimed or intended.

Recommendation:

Support the development of standards for validation of SP tools. This will be strongly application dependent, but good ideas will benefit several applications.

Human Interface

Finding:

Existing human interfaces to hypertext (the format used in the World Wide Web consists of mixed text, images, and sound with "hot words" to be clicked to produce more information) and multimedia on the NII are inadequate and require improved speed, efficiency, and ease of use.

Recommendation:

Support research on SP methods to improve the interactive use of hypertext and multimedia information on the NII. Develop techniques for the automatic integration of information into databases and the means for easily extracting the information (as detailed in Section 5.10.)

Interactive Teaching

Finding:

Interactive teaching has been developed very rapidly during the past few years, especially as it has concerned simple textual data. Interactive teaching should include multimedia such as voice, pictures, text, video, and graphics. Separate educational disciplines are independently developing their own methods and teaching files. It would facilitate development of the general field if common methods and shared applications were emphasized.

Recommendation:

Support research aimed at developing common SP tools for multimedia interactive teaching and general database architectures for implementing interactive teaching files. This will require cooperative efforts among the SP, database, and education communities. In particular, it will require the active cooperation of the users in the design of the tools.

Public Education about NII

Finding:

The NII exacerbates the problem of "what is real," because it brings to users lots of questionable data — digital images and sounds. Yet, the ground truth or gold standard for these things is far away, essentially inaccessible except through the easy-to-manipulate digital medium. An understanding of the fundamentals of SP will enable users to better understand the potential for both accuracy and deception in digital representations.

Recommendation:

Fund exploratory R&D targeted at bringing SP concepts to the general public, especially through K-12 education.

1.4 Conclusions

The primary goal of this report is to inform both research agencies and the general NII community of the key role of SP in the NII and to identify important trends in the development of SP for the NII. We believe that active cooperation among the SP community and many of the user communities of the NII will provide the fastest possible development of new and widely useful SP tools for the NII. The National Science Foundation and other Government agencies can best assist such development by encouraging the necessary interdisciplinary research and development efforts by whatever means possible. This report describes several possibilities.

Chapter 2

Findings and Recommendations

2.1 Introduction

By one measure the NII will be a success if it can provide wide distribution of education, information, and entertainment services to the community as a whole. An example would be the delivery of educational materials to local schools, colleges and homes. The material will most likely be interactive, and it could well involve tutorial games. Other examples would be the remote practice of medicine directly in the home for routine problems, or telecommuting. These applications have in common that they will be more effective if provided over channels that allow the transmission of multimedia signals rich in video data and voice. On one hand, these signals require a significant bandwidth (1 Mbit/s or more,) and on the other should be distributed rather broadly, ideally to every home.

While existing telephone network can meet the ubiquity requirement, bandwidth presents a problem. One approach is to use complex modulation schemes to transmit the information over existing (or slightly upgraded) copper lines. An alternative approach that is gaining popularity is to use a mixture of fiber and cable. In either case, channel capacity is costly to provide and signals need to be coded and packaged efficiently for transmission to the office, school and home. These signals must then be individually processed at each terminal or station, adding significantly to the potential cost of the station.

And here is the challenge if ubiquity of service is to be achieved without heavy subsidy: the cost of each station must be modest. Today a cable-TV set-top box costs less than $200 to manufacture; perhaps a box providing community information services could cost as much as $500 to manufacture (and sell for less than $1000) and still reach high penetration. Contrast this with terminal equipment that Time-Warner are using in the Full Service Network trials in Orlando Florida, where the equipment cost is reputed to be *many* thousands of dollars.

The terminal equipment performs a lot of very sophisticated SP. It includes the demodulation and recovery of the digital signal, protocol processing to recover the information streams, decoding to obtain the content of the streams, and frequently sophisticated rendering and synchronization in presenting the information to the user. And this just covers the one-way stream from the supplier to the user. The complexity of the return stream will depend greatly on the application. It could vary from little more than a low-speed data channel for browsing an information base, to a real-time video channel if video presence is required, as when a telecommuting manager tries to assess the status of a project.

Many forecasters in the business of community services expect the cost of the terminal equipment to fall significantly over the next few years. However, the sophistication of terminal functions is such that a reduction in cost will not occur unless significant advances are made in a number of technology areas, perhaps the most significant being SP algorithms. Findings and recommendations for achieving high performance at low cost are presented in this chapter. The interested reader can find an ordered tour through the state of the art as well as probable and desirable directions for the future in Chapters 4 and 5 which provides the background and motivation for the topics of this chapter.

2.2 Grand Challenges

Universal Access for Education

Although the phrase "universal access" has lost some of its significance by becoming media buzz words, the goal remains clear and important: the development of SP for the NII should incorporate the goal of providing useful information and applications to all users, including those with minimal facilities and training. An example of how SP can contribute to this goal is the method of progressive and multiresolution transmission, where a single provider of images with a single image compression algorithm can provide adequate pictures to "bit-starved" users such as the antiquated computers in an elementary schoolroom connected to NII by relatively slow modems and telephone lines, while still providing high quality pictures to those with more sophisticated equipment and needs. The former have access to fewer bits but get the best picture they can within those constraints, the latter efficiently use the additional bits to improve the picture to the high quality needed for scientific research or medical diagnosis. To support this goal we must develop more efficient, low cost, reliable techniques for the acquisition, storage, transmission, and display or reproduction of information for the wide variety of subnetworks and computing equipment found on the NII. Of particular relevance to this goal are SP algorithms which are "scalable" in the sense that signals on the network can be reconstructed at a variety of quality levels depending on the equipment available to the user. Many of the specific recommendations to follow are of particular importance to promoting wide access because of their goal to improve and simplify the human interface with NII and to make it useful to the widest possible audience.

Tools to Support Content Analysis

We are rapidly approaching an "information watershed" wherein the challenges shift from generating, storing and transmitting data to problems of integrating, digesting and comprehending information. A line is crossed when affordable and ubiquitous electro-optical bandwidth exceeds the bandwidth of the human nervous system that is devoted to sensory input and motor control (visual, aural, tactile, etc.). With terabit-per-second optical links now being tested, this watershed event is not far off. When communications bandwidth approaches the human processing limit, further increases in data rates make sense only if they are balanced by proportional increases in automatic integration, filtering and presentation of that data, plus the means to browse through enormous databases to locate data with content relevant to the user's interest.

Health care is an excellent example. The combined totality of patient records, medical literature, insurance guidelines and regulations, pharmaceutical data, and other information already overwhelms the typical practitioner. It has been reported that many physicians now spend over 40% of their professional time on paperwork associated with record keeping and billing. Delivering 100% more raw data to this community makes little sense unless accompanied by value-added preprocessing power.

This observation provides a strong argument for more aggressive interdisciplinary research in methods that extract needed content from signals by performing modeling, identification, classification, segmentation, recognition, and other forms of data analysis. These challenges can only be met by researchers capable of spanning a broad intellectual and scientific range that includes SP, applied mathematics, computer science, statistics, linguistics, and systems/control theory.

Data Quality Validation

SP changes data. The intent is to make the data better in some way, such as more efficient to communicate and store, faster to display, or easier to interpret by a human observer. This raises the important issue of data quality and integrity: how does one validate SP algorithms for a particular application in the sense of quantitatively demonstrating the quality and utility of their output? These issues are particularly important in medicine and science where "enhancement" to one user might be degradation to another, or the "noise" removed for one user might be the cherished signal to another. It is critically important to develop acceptable and trustworthy protocols for validating the utility and quality of signals altered by SP and to develop effective means of predicting the utility and quality from easily made quantitative measurements. This can involve simulations of user practice, statistical analyses of differences in decisions or measurements made on processed versus unprocessed data, human perceptual models, subjective testing, and the development of improved quality metrics for signals that are seen and heard.

2.3 Cooperative Research in Key Areas

Finding:

SP research goals common to a variety of applications can be identified, but theory and algorithms must be developed in the context of intended applications in close cooperation with users if the algorithms are truly to reflect the needs of the user community while having the highest possible efficiency and performance. This demands interdisciplinary research, often involving many organizations and institutions as well as researchers of widely varying backgrounds. Such research can be extremely difficult to conduct because of institutional barriers, the lack of a single clear source of support, and the cultural differences among the various technical fields and user communities.

Recommendation:

Genuine, interdisciplinary and cooperative efforts for research on SP for the NII among academia, government, and industry should be initiated by every party and supported to the greatest extent possible. To improve existing SP tools and develop new ones we need to:

- Improve essential tools for acquisition, decomposition, compression, transmission, reconstruction, modulation, coding, equalization, and other SP techniques. Develop new and better standards and encourage promising novel approaches. It is important to tilt the balance of support toward the very best research rather than incremental improvements of mature methods or unsubstantiated claims of allegedly novel methods.
- Emphasize scalable/progressive/multiresolution approaches in signal decomposition, compression, and other SP.
- Develop asymmetric image processing and transmission — for example, modest computing power for the student in a classroom, extensive computing power for the imaging center.
- Develop fusion tools to combine data of different types, especially imagery.
- Develop information extraction tools for browsing and data extraction.
- Develop tools for automatic multimedia database creation, management, review, validation, indexing, abstracting.
- Develop perceptual coding methods of speech, audio, image, and video. This likely necessitates improved models of human perception and quantitative signal quality metrics.
- Develop efficient models of human speech production, images, including the “talking head” or “wire frame” models for very low rate video telephone.
- Develop credible validation procedures for quality and utility of compressed and otherwise processed signals in comparison to originals or other benchmarks.
- Develop algorithms for segmentation, classification, transformation, recognition, and registration for speech, images, and other signals. Develop efficient algorithms combining these SP types with each other and with other SP such as compression and error control.
- Develop information agents for the automatic searching of large databases by content.

There are a variety of means by which the development of SP tools can be encouraged within the context of specific applications through interdisciplinary research.

- Support collaborative research in critical areas among universities, industry, government, and users. The government can evaluate proposals to determine scientific merit as well as national priority as a basis to provide funding. The government can select specific topics such as those generated by this workshop and encourage and sponsor collaborative research by academia and industry. Genuinely interdisciplinary and collaborative research requires cooperation among Programs, Divisions, and Directorates within an agency and among different agencies. Target areas might be selected for joint review and funding. Many governmental agencies (NSF, ARPA, NIH, NLM, NIST, NTIA, NEA, NEH, etc.) issue extramural research grants and contracts through various mechanisms, including the extramural grants activities and broad agency announcements. In addition, individual experts can be supported by purchase order, intergovernmental personnel act (IPA) appointments, or as visiting scientists. A more recent approach is the Cooperative Research and Development Agreement (CRADA) which is a means to transfer technology from federal labs to the private sector, allowing private groups to take prototype designs from the labs and develop them into commercial products, sharing any ensuing royalties with the government. Such programs explicitly targeted at SP for specific NII applications should be generated and widely announced. The NII can itself be useful for communicating opportunities, possibly via a newsgroup.

- Use the NII to foster the NII. At this time useful data exist on the Internet, but they are often widely dispersed and sometimes hard to find. The NSF could establish a focal point and then moderate the discussion to encourage interaction in line with its goals. Integrating this into government research funding would give it real focus (e.g., if the research grant development and approval process required use of these resources). This could be accomplished by managing an Internet news group specifically geared toward SP for the NII.

- Fostering interdisciplinary interactions is difficult, it takes time and money and people to push it. We might try focused workshops of invitees from the participant communities organized around tutorial presentations, held year after year with a core group of repeat participants. Graduate student participation is vital, to populate the interface. Post-doctoral fellowships which send students from one community to another for a few years are also an approach. Money is more important as a lure in interdisciplinary work than in more established areas where people end up anyway. Agencies such as the NSF with a strict discipline orientation may need to alter their grant management structure to provide adequate time and program management continuity to foster broad interdisciplinary research.

- In view of the centrality of SP research for the NII to future technological developments, it is particularly important to facilitate the participation of members of historically underrepresented groups in this area. Thus we suggest that proposals be actively encouraged and supported to encourage workshops for women or minorities, to be appended to the annual SP symposia, and post-doctoral fellowships so that women or minorities from areas such as physics, statistics, or math can participate in SP projects.

2.4 Standard Databases

Finding:

Theoretical development of SP algorithms can often take place using easily available data for preliminary tests, but genuine comparisons among competing algorithms and validation of quality and utility require generally available standard data sets. For example, computer aided diagnosis of CT scans or feature extraction from telemetry data for topological analysis may behave quite differently on some data sets than on others. Hence common standard databases can enhance research and ease comparison among existing schemes and competing new schemes.

Recommendation:

Support the establishment of standard databases for SP research. This can be done either through financial support to individual institutions (as may be necessary for expensive databases such as high resolution digital mammograms) or by providing a clearing house of publicly available databases, reachable via the NII. The NHANES database of digitized lumbar and cervical x-rays created by a collaborative project among the National Library of Medicine and two other federal agencies provides a prototype. The existence of these databases and instructions for their use should be made widely available to relevant NII newsgroups and professional newsletters.

2.5 Standards for Tool Validation

Finding:

Not all new SP algorithm developed over time can or should be made readily available for NII use. Many have become and will become *de facto* standards simply because they work well and no clearly superior algorithms performing similar or better functions have appeared. Nonetheless, it will be useful to develop standards for the validation of new algorithms, that is, to demonstrate quantitatively that the algorithms perform as claimed or intended.

Recommendation:

Support the development of standards for validation of tools. This will be strongly application dependent, but good ideas may prove useful in many applications. Examples include quantitative methods for fair comparison of compression methods in terms of quality, bit rate, speed, and cost; or methods for verifying diagnostic or classification accuracy in computer processed images in medicine and science.

2.6 Human Interface

Finding:

Existing human interfaces to hypertext (the format used in the World Wide Web combining text, images, and sound) and multimedia on the NII are inadequate and require improved speed, efficiency, and ease of use.

Recommendation:

Support research in SP methods to improve the interactive use of hypertext and multimedia information on the NII. This includes the development of techniques for the automatic integration of information into databases and the means for easily extracting the information as detailed in Section 5.10. There should be NII support for communication between the SP developers and users. For example, automated indexing algorithms should be developed in conjunction with indexing specialists. Like librarians who catalog books, NII indexing specialists will be needed to do classifications, develop indexes, etc. Communication "institutions" should be established so that these people all talk to each other as they do their work, to avoid the kinds of cataloging idiosyncrasies that have plagued the book library system.

2.7 Interactive Teaching

Finding:

Interactive teaching has been developed very rapidly during the past few years, especially as it has concerned simple text. Interactive teaching should include multimedia such as voice, pictures, text, video, and graphics. In the health-care delivery field, each subspecialty is already developing teaching files of its own. We must exploit the commonality of these databases to facilitate the development of interactive teaching.

Recommendation:

Support research aimed at developing common SP tools for multimedia interactive teaching and general database architectures for implementing interactive teaching files. This will require cooperative efforts between the SP, database, and education communities. It should actively involve the cooperation of the users of the tools.

2.8 Public Education about NII

Finding:

The NII exacerbates the problem of what is real, because it brings to users lots of questionable data — especially in the form of digital images and sounds. Images and sounds can be artificially generated or altered from their original form. Means of validating these signals may be remote, essentially inaccessible except through the mistrusted digital medium itself. Being able to assess "what is real" or genuine or true will be a basic part of understanding the information available to us.

Recommendation:

Fund exploratory R&D targeted at bringing SP concepts to the public, emphasizing general public education and inclusion in the K-12 curriculum. This must involve cooperative efforts between the SP and education communities.

Chapter 3

Introduction

3.1 The National Information Infrastructure

During recent months there has been an explosion in the media and in public awareness of the National Information Infrastructure (NII) or "information superhighway" and its most famous component, the Internet. From *Time Magazine* to *Technology Review*, from trade journals to daily newspapers, from alumni newsletters to professional magazines, the NII has moved from the Science and Technology sections to the front page, with regular stories about its history and nature and its numerous applications in education, commerce, medicine, government, and science. Stories abound on the many new electronic locales in "hyperspace" for groups sharing common interests, and of new public access to high-tech marvels such as the images of the recent comet collisions with Jupiter being acquired by NASA, transmitted, stored, displayed, and studied by grade school students. The media blitz is puzzling to much of the technical community, who have actively used the Internet and its many powerful tools for more than a decade almost without conscious thought: telnet, ftp, email, and bulletin boards have been in daily use for several years by most scientists and engineers as well as many in the library and medical communities. The ability to share data, acquire literature, and stay in close and constant touch with colleagues around the world is taken for granted in the technical community. The recent discovery of the NII by a much wider audience and its promotion as a national effort with profound economic and political as well as technical implications is bound to have a fundamental impact on its future development and use. As a wider public gains access to the most popular tools of the NII such as telnet, ftp, archie, gopher, veronica, wais, the World Wide Web (WWW), and mosaic, the great potential for public benefit is threatened by the perhaps greater potential for chaos due to the sheer volume of users, nodes, applications, tools, and networks. The blessings of wide access to library catalogs, software, educational information, interest groups, travel information, and consumer products are dimmed if the sites sought are too crowded to enter, if the image data take hours to download, if the data are inaccurate, if sensitive files are read or altered by unauthorized entry, if the existing tools do not enable the user to find the information sought in a timely fashion, or if the data are incompatible with the applications software.

Given the *laissez faire* principles underlying much of the explosive growth of the NII with its minimalist approach to central control, it is critically important that the principles enabling the net be understood at least in general terms by those who influence its development, and that the technical community actively participate both in educating the public and in the development of the networks. Conversely, it is absolutely essential that portions of the user community cooperate with those developing the tools in order to ensure that the tools are genuinely useful in real applications. In any large scale system like the NII there are delicate balances among competing goals. There is a constant tension between the standardization necessary for compatibility of the many components and the evolving technologies and software that promise better performance, between maintaining the integrity and security of the data while making it available to the desired community, between permitting the widest possible access and maintaining network speed and reliability, between preserving an enormously valuable scientific and technological tool for research and development and providing consumer services for profit.

The success and utility of the NII Network will depend on its ease of use and the number of tasks it can perform. For this reason the emphasis must be on making the NII user-driven and user-friendly. The NII must be a transparent and seamless network that is capable of accommodating users with minimum facilities.

3.2 Signal Processing in the NII

The NII is built on a technical foundation whose building blocks straddle many fields. It contains many layers of both hardware and software and few (if any) individuals are expert in the full array. This report aims at one enabling technical area — SP — and at several specific applications where SP is crucial for future development. The field of SP consists of the theory and practice of software and hardware that convert signals such as speech, images, video, sensor data, music, and EKGs into a form useful for a specific purpose, such as transmission, storage, display, diagnosis, and interpretation. The software for a particular SP technique implements an *algorithm*, which can be viewed as a recipe for converting an input signal such as an image produced by a CT scanner into an output such as a computer file suitable for efficient and reliable transmission through a network. These processes serve a variety of specific applications in education, government, medicine, geographical information systems, research in many scientific and technical fields, entertainment, and simple satisfaction of curiosity. The large majority of practice concerns digital signal processing (DSP) using computers or special purpose digital hardware. Perhaps the most common SP in the NII is the repertoire of methods for the efficient and reliable transmission and storage of data. SP also occurs in many applications particular to the data, including the processing of speech and images for enhancement and recognition, and statistical feature extraction and pattern recognition used to extract the important or relevant information from large quantities of accessible data.

The NII exists today partially because of recent developments in the theory and practice of several basic areas related to SP. The basic networks can be viewed as a combination of network protocols such as the internet protocol(TCP/IP, not here considered to be within the domain of SP) and techniques for the reliable and efficient acquisition, storage, communication, interpretation, and display of information. The latter extensively use DSP such as data compression for speeding transmission and minimizing storage requirements, error control coding for accurately reconstructing information following errors in storage or transmission, modulation for efficient communication over digital and analog links, equalization for combatting distortion arising on real channels, and a variety of linear and nonlinear filtering methods for improving perceived signal quality. As the NII develops, SP is responsible for the improvement of current methods and the development of new tools and applications.

3.3 Standards

The consolidation of a major component like NII into the nation's infrastructure creates the opportunity for defining standards and for establishing protocols for defining standards. This process is especially appropriate when the nation itself is the sponsor of NII, and when national prosperity and security depend upon it. Examples, past and current, include radio-band allocation and modulation standards, highway specifications, power distribution, and data-security systems.

Such deliberations marked the choice in 1976 of a national cryptographic standard to ensure privacy and authenticity in digital network communications. The Data Encryption Standard, promulgated by the NIST (formerly the National Bureau of Standards), has been in successful use ever since. It is important for our purpose to observe that hardly any objection was raised to the rationale for a standard or to its establishment, but rather to alleged weaknesses and to the exclusion of the public from the process of choice. In the same way, we may expect the public to welcome standards for the NII, so long as they are clear, well justified, and as non-restrictive as possible.

In its request for proposed standards in the May 15, 1973 Federal Register, the NIST specified these criteria (among others,) which might well describe our own recommendations: An acceptable algorithm must be

1. completely specified and easy to understand;

2. available to all users;

3. adaptable for use in diverse applications;

4. economically implementable in electronic devices and software;
5. efficient to use;
6. able to be validated;
7. exportable.

3.4 The Panel and the Report

In August 1994 a workshop/panel of experts (listed in the Appendix) in various areas within or related to the field of SP was convened for two days at the National Science Foundation for the purpose of discussing the role of SP in the development and use of the NII and speculating on specific topics particularly important for future growth and applications. This report summarizes those panel discussions and the interaction taking place by email prior to and following the meeting. The discussions focused both on the development of enabling technologies — the underlying tools that are useful for a wide variety of applications — and on several specific applications with strong SP components. The enabling technologies include efficient methods for coding, compressing, encrypting, communicating, storing, searching, interpreting, and displaying image, audio, and other data. The applications include medical image processing, geographical information systems, consumer multimedia, remote sensing and planetary research, and education.

Chapter 4

Signal Processing in the NII: Past and Present

4.1 Introduction

From the middle of the nineteenth to the middle of the twentieth century, manufacturing capability was closely coupled to economic development and wealth. The last third of the twentieth century is often termed an "information age," and information capability has served as the source of great surges in economic development. For our nation, the American communications infrastructure is the essential enabler of this information age. Less recognized is the critical role played by SP in the practical development of this communications infrastructure, and probably even more so in the development of related commercial applications. The American communications infrastructure can be likened to the power system infrastructure established primarily during the first two-thirds of the twentieth century. This power system has provided far more than just the ready availability of electric light to the individual home. It is the backbone of our manufacturing capability; it provides the basic electric needs of homes, businesses, schools, and hospitals. The information infrastructure is rapidly evolving to a similar utility status. It is the essential enabler of numerous applications that will form the basis of significant economic development well into the next century. SP is the critical technology to make many of these applications possible.

Digital SP (DSP) is the theory and practice of the representation of information-bearing waveforms as discrete signals, and the digital processing by computer or digital circuitry of such signals in order to extract information or accomplish a variety of tasks. The dominance of digital over analogue computers has resulted in the vast majority of SP being done digitally, but sensory data are inherently analog or continuous, so even DSP systems often *begin* with continuous SP.

The origins of SP lie in mathematics developed over the last three hundred years, but the development of modern SP has been in close linkage to digital communications, digital control, and digital computers. Indeed, it can fairly be argued that without SP enabling modern digital communications, the NII would not be possible. The acquisition, conversion, coding, transformation, storage, transmission, reconstruction, display, and interpretation of data or signals all involve SP. The data might be voice, audio, medical or scientific images, remote sensors, satellites, telephones, computer applications, or the stock market, but within computers and networks all such signals are represented as numbers, and most transformations from one form to another can be viewed as SP.

One salient feature of research in SP is the relatively rapid transfer of theory to practice. For example, within one decade of the introduction of coded modulation, (error-control coding combined with efficient modulation) and of adaptive equalization methods, the available data speeds for low cost modems shot from 2,400 bits per second (bps) to 28,000 bps. Linear-prediction theory and vector quantization have led to international standards in speech compression systems providing high quality at only 4,800 bps instead of the former 64,000 bps required by uncompressed digital speech. Advances in digital-to-analog conversion combined with error control coding made possible the explosion of digital music on CDs in consumer electronics.

4.2 State of the Art

SP encompasses a variety of areas of application and basic research. A review of some of these areas emphasizes their critical role in the development of the NII and in the SP tools taken for granted in many of the user applications on the NII.

Signal Acquisition and Synthesis

Signals arising from natural and artificial sources such as speech, music, sensor readings, and images must be captured and converted into a form that computers and other DSP circuits can handle. Following processes such as storage, transmission, or enhancement, some form of the signal might be reconstructed from the computer representation. Data acquisition involves tools such as analog-to-digital (A/D) conversion, which itself can involve sampling — selecting only certain values — and quantization — approximating by rounding off continuous phenomena to computer-compatible discrete representations. Data synthesis often involves the reverse operations: digital-to-analog conversions to produce signals that can drive a physical "output" device such as a loudspeaker, a terminal display, a television, or an antenna. During recent years technology has continually improved on the quality of approximation and the cost of production of a variety of converters between the analog and digital worlds, and such devices now exist in most personal computers as well as in commercial high resolution scanners and high fidelity audio.

Signal Decomposition

A common theme in many SP applications is to decompose an original signal into its primitive or fundamental constituents and to perform simple operations separately on each component, thereby accomplishing extremely sophisticated operations by a combination of individually simple operations. The classic and still pervasive example is Fourier analysis, the theory and practice that breaks signals into sinusoidal (smooth oscillating) components just as complex sounds can be decomposed into rich combinations of simple tones. Fourier analysis lies at the heart of the development during World War II of modern analog communications and radar technology, and discrete versions play a corresponding role in DSP. Such decompositions have been a key theoretical tool for the modern analog and digital communications techniques that underly transmitting signals through the physical pathways of the NII. They have also proved equally important to the understanding and advancement of error-control coding for reliable communications, SP for removing unwanted noise and signals, and SP for improving the quality and utility of signals. During recent years Fourier methods have been supplemented by other approaches, most notably the many methods now subsumed under the general heading of *wavelets.* These alternatives hold promise for providing more useful ways of analyzing and processing signals for different applications, but the classic methods still provide a useful theoretical tool for evaluating the new tools, and continue to dominate in applications.

Signal compression

Signals such as speech and image data are stored and processed in computers as files or collections of bits. Big files have many disadvantages in comparison with small files: they fill storage devices, they take longer to transmit to users, and they can overwhelm algorithms designed to draw conclusions from the raw data. Anyone who has patiently waited for an image to be transmitted through a modem appreciates the value of a speedup of 10 times, or even 2 times. It is clearly desirable and sometimes necessary to reduce the number of bits required to describe a signal, but it is equally clear that this must be done with great care if the reduction results in a loss of information. Signal compression or data compression is concerned with reducing the number of bits required to describe a signal to a prescribed accuracy. Because compression is often one of the first processes applied to a digital signal produced by a source of information such as a microphone, sensor, or camera, the combination of quantization and compression is called *source coding* in the theoretical literature, and the application to speech and image data are called speech coding and image coding, respectively.

Signal compression is a key technology for multimedia applications on the NII. Many current applications treat sound, images and even video as uncompressed bitstreams, but this situation will change when their use becomes widespread. Storage and bandwidth limitations call for powerful compression methods.

Compression algorithms fall into two general classes:

- *Lossless Compression* A file is compressed in such a way that the original file can be perfectly recovered from the compressed file. The Morse code of telegraphy provides a classic example: by sending short patterns (dots and dashes) for frequent letters, and long patterns for rare letters, fewer bits are needed on the average than with the standard computer representation (ASCII,) which assigns the same number of bits to all letters. Most popular algorithms nowadays are based on adaptive Huffman or arithmetic coding, or Lempel-Ziv (LZ) methods.

- *Lossy Compression* Compression is lossy if the original file cannot be perfectly recovered from the compressed file; the goal is to get the best possible quality for the given compression. *Quantization itself is a form of lossy compression.* Any compression on a continuous signal (like speech or an x-ray) is necessarily lossy. Most popular lossy compression algorithms perform some form of transform or predictive coding, where the signal is first decomposed and then quantized.

The main shortcoming of lossless compression is that the amount of compression is limited, with typical compression on computer data files being about 2:1. Commercial, freeware, and shareware programs for lossless compression abound throughout the NII; they usually succeed in halving memory requirements for text and program files.

Quantization causes loss of information, even if that loss is not perceptible. Quantization can occur in the original A/D conversion, in the digital arithmetic used to transform a signal, or in quantization mapping a digital signal with a high bit rate into a compressed digital signal with a low bit rate. Such quantization can operate either on individual samples or pixels (scalar quantization,) or on groups or blocks or vectors of samples or pixels (vector quantization). In his pioneering development of the mathematical theory of communication, Shannon showed that that better performance can be achieved by using vector quantization, but scalar quantization is simpler and often adequate. The international JPEG (Joint Photographic Experts Group) standard achieves compression of digitized images by 16:1 with reasonable quality using scalar quantization.

Image compression is already widely available on small computers through compressed versions of standard image formats like TIFF and GIF as well as the JPEG standard and Apple Quicktime. The JPEG standard for still imagery uses a discrete cosine transform (DCT) scheme with scalar quantization combined with lossless coding and achieves reasonable quality with 16:1 compression. Software implementations are widely available and existing special purpose chip implementing this technique operate at 10 Mbps. Other related standards such as the H.261 standard (popularly referred to as p$\times$64) and and MPEG 2 are now well established, but current standards committees developing future standards such as MPEG 4 for low rate video have not yet agreed on which compression approach to adopt. The MPEG 2 standard for video compression is the basis for the U.S. HDTV broadcast standard currently being tested by the FCC in conjunction with the AC-3 multichannel audio compression algorithm.

Proprietary vector quantization algorithms provide fast decompression of video even on personal computers, and are included in such popular software as Apple QuickTime and IBM Photomotion. The current code-excited linear-predictive (CELP) international speech compression (or speech coding) standards combine sophisticated SP with vector quantization, and accomplish compression ratios of about 12:1. Speech coding has enjoyed a recent rebirth in importance because of the growth of telecommunications networks, cellular telephony, and personal communications systems. The coming worldwide mobile telephony systems will add even more importance to this technology. Speech coding also plays a critical role in voice mail and satellite paging systems with voice messages downloaded to pagers. Speech coding is moving into personal computers and hence to immediate NII access. For example, Microsoft's new Windows release (nicknamed *Chicago*) will have speech coding feature called TrueSpeech.

Lossy compression algorithms are important for the rapid transmission of voice and image data across the NII and are widely available. They are often components of far more complex algorithms such as Mosaic, the increasingly popular software for finding text, sound, images, and video in the World Wide Web through the NII. By providing scalable transmission, they permit the user to balance signal qualities against data rate. Without such flexibility, browsing is not possible, and massive archives lose their utility.

Existing popular compression algorithms are based on ideas that have matured over the last 10–15 years in academia and industry. During recent years the emphasis in the compression community has shifted from that of developing new techniques to that of software and hardware implementation of time-proven techniques, and the development and establishment of standards to achieve interoperability of equipment. Of course, as existing systems improve and new compression systems are devised, these standards will be modified and updated. At the same time, choosing these standards will have a major impact on future developments and growth of various bandwidth compression techniques. These standards have already had a tremendous influence on the "digitalization" of the consumer electronic world, and they have led to digital HDTV where the U.S. currently leads. In high quality audio coding (CD-quality digital audio compression), advances in filter banks, perceptual coding and masking have led to standards such as MUSICAM, which in turn will permit digital audio broadcast to replace FM radio. The technology of high-speed and high-resolution analog/digital conversion, especially in digital implementations, has been instrumental in all applications whose data originates in the real world.

Progressive and Multiresolution Signal Transmission and Representation

In many applications of quantization, signal decomposition, and compression it is useful to be able to describe, transmit, store, or reconstruct a signal at different scales, resolutions, or quality levels. The idea is exemplified by having a single data source of bits on the network serve both the modestly equipped user with obsolete computer on a slow phone line in an elementary classroom and the high-end user doing industrial design or scientific research with a new workstation, high resolution monitor, and NII access through high-speed links. The low-end user could extract a postage stamp size reproduction of reasonable quality, while the high-end user could produce a full screen version of sufficient quality for critical judgements. A key in distributed applications and Internetworking is the access to data at various quality levels that depend on the available bandwidth or terminal capabilities. This is achieved with scalable (multiresolution) and progressive coding: coding methods that allow a common data stream to yield signals of varying quality to diverse users.

Digital Filtering

Once a signal is in the digital form compatible with computers, it may be altered using digital filters. These are reasonably simple algorithms that can be implemented in software or hardware and can change a signal for a variety of purposes, including shaping it into a form suitable for transmission or storage, removing noise and distortion, sharpening edges in images, or minimizing interference with other signals. Such filters appear in virtually all components of communications systems and may be linear or nonlinear.

Modulation

Modulation and demodulation convert the signal to a form appropriate for transmission through optical fiber, coax cable, wireless media, phone lines, etc., and reconvert the signal upon arrival. These schemes can employ classical radio methods (analog modulation like commercial AM and FM) through modern methods based on statistical SP using matched filters and correlation detectors and maximum likelihood demodulation which are geared specifically to the transmission of digital signals (digital modulation). The theory and practice of modulation is mature and well understood, but new applications and devices add constant pressure for new extensions and directions.

Adaptive Signal Processing

Many SP algorithms adapt to changes in their environment by constantly measuring their own performance and adjusting their own parameters. A particularly important example is adaptive equalization by a receiver to determine and remove the time varying behavior of the channel through which the transmission takes place. For example, different cables in the telephone network, the changing nature of the path between a mobile cellular phone and the nearest repeater, or varying atmospheric conditions around a satellite dish can drastically alter a received signal.

Error Control Coding

Error control coding incorporates information into the signal that allows a receiver to find and correct bit errors occurring in transmission and storage. Since such coding detects or corrects errors incurred in the communication or storage channel, it is often referred to as *channel coding*.

Recent advances in SP have allowed us to achieve performance on certain communication channels near the theoretical limit characterized by information theory. These successes were attained using advances in adaptive channel equalization, error correction coding, modulation techniques, and joint coding/modulation such as trellis coded modulation, where the error control coding is combined with efficient modulation techniques. High speed transmission over the local telephone loop and hence the potential for universal high-bandwidth access to the NII from homes depends on such high speed SP. This again shows the power of SP theory, algorithms, and technology in solving a real need for the NII.

Channel coding for fixed channels is a relatively mature field, but coding for time varying channels is a relatively young and immature field. Such channels are increasingly important, as exemplified by the wireless channel used for personal communication systems which are beginning to access the NII.

Channel coding for storage has also been very successful in boosting capacity, and remains a topic of active research. Much of the error control necessary for CDROM capacities was adapted from space communications technology, and the coding for rewritable CD's advances the SP devised for magnetic recording.

Classification and Recognition

All of the methods considered so far essentially try to reproduce the original waveform provided to the SP algorithm with the best possible fidelity; they make no pretense of extracting content from the signal. Algorithms for extracting features or patterns from raw data and reporting them in a useful way have long existed and are the subject of active research, but they have not yet developed into stable tools like those mentioned above. They have thus far had a limited impact on the NII, but this is likely to change rapidly. Such "information value-added" algorithms form the core of the future directions described in Chapter 5. Simple variations on these ideas have already been seen in earlier algorithms: for example, deducing whether a signal received over a noisy channel represents 0 or 1 constitutes a form of pattern recognition. But much more elaborate schemes currently exist or are under development.

- Speech recognition: Methods exist for converting spoken speech acquired by microphone into text suitable for printing or for giving commands to computers. Machinists can instruct robots in noisy environments on assembly lines, radiologists can dictate reports during screening which are automatically transcribed. These activities could easily take place over the NII, for example, for remote control of scientific experiments, or control of robots in hazardous environments, or delivery of diagnoses to a remote clinic from an imaging center.

- Image recognition and segmentation: Methods exist for automatically breaking up images into interesting parts to assist human interpretation or to allow computer tracking of changes. The speed and accuracy of these methods depends strongly on the particular applications, but they show promise for such applications as automatic environmental tracking and computer-assisted diagnoses.

- Optical character recognition (OCR): Most scanners for personal computers now come with OCR software which (to some degree of success) can scan a page of text as an image, interpret the letters, and save the information as a text file rather than as a "photograph" of the page. Such programs can play a crucial role in making old books available to NII users by creating electronic databases in an almost automatic fashion.

- Pattern recognition: Changes in patient state are often indicated by well known correlated changes in blood pressure, pulse, EKG, etc., that can be detected by computer. Early detection alerts local and remote personnel of a dangerous turn of events.

Algorithms of these types have great potential in making the NII more useful. For example, enormous databases could be searched for all examples containing prescribed features — think of searching recorded speeches for all excerpts containing the words "information superhighway."

Imagery Data Fusion

In most practical applications the combined information of two or more sources of complementary data can be greater than the sum of the parts. This is found to be true when imagery data needs to be annotated for location identification, or when maps and imagery need to be overlaid for better identification of lines of communications. Indeed it has been shown that the performance of automatic machine exploitation and identification techniques improves dramatically when complementary data is combined and is used as the input to the exploitation device.

There are several current imagery-fusion systems that can be used as initial models for NII; consider these:

1. The commercial ARC/INFO fusion system combines geographic data with attributes to represent complex spatial relationships for easier decision making. This system reformats the input spatial data to the area of coverage, combines it with the attribute data, projects the combined data to the geographic database to form a new data relationships. An example is combining elevation data with a weather map and Landsat imagery to form an image that facilitates decision making for a commercial pilot and also aids the automatic landing system.

2. NASA EOS Data and Information System which is being developed to manage data resulting from NASA's Earth science research satellites and field measurements programs. EOSDIS will generate user-defined data products and will facilitate the combination and manipulation of data from all sources. This is an example of a user-driven distributed system that must be interconnected using a diversified set of links to function as a seamless network.

3. Commercial image processing tools for fusion of imagery data such as ERDAS, MCIDAS, ENVIk.

These information and imagery-fusion systems are in early stages of development and require significant improvement to become truly automatic real-time devices. The areas of SP that must be developed to make this a reality are discussed in Sections 5.8 and 6.2.

4.3 Standards

Standardization activities tend to have a major impact on research directions, and reciprocally, research has had a major influence on standards. One dramatic illustration: in the area of speech coding (and possibly also with MPEG-4 and MPEG-NewDirection,) standards groups have established performance requirements and objectives (so-called "terms of reference") for a future standard that *far* exceeds the state of the art at the time. Thus, major challenges confront researchers to achieve these objectives if they are indeed feasible. The speech-coding group, seemingly unacquainted with fundamental performance bounds imposed by the distortion-rate limits of Shannon theory, just extrapolate from prior progress and assume that in two or three years present quality will be available with a further drop in rate. Amazingly, thus far this strategy has worked. An example is the recently adopted CCITT (now ITU-T) recommendation G.728 for 16 kb/s speech coding, a universal standard for telecommunications. When the terms of reference were established in 1988, they appeared to be impossible to achieve; likewise with the ITU-T 8 kb/s standard, whose terms of reference are now nearly attained.

These are cases where innovative research is driven by well-defined and challenging targets. Much of this directed research has been taking place in industry. On the other hand, standards have been informed by the availability of a body of past research, often without directed applications in mind. For example, the G.728 standard uses gain-adaptive vector quantization and pseudo-Gray coding, which originated from university research.

4.4 Applications

We have touched on some of the most important, well established and mature SP methods appropriate to the NII, but a wide variety of special techniques can also be found in particular applications. Common examples include motion compensation in video compression, where automatic extraction of motion in an image can permit significantly improved compression, and related deblurring algorithms for improving still images taken by or of moving objects. Worthy of mention also are synchronization methods, noise removal and enhancement techniques for improving the perceived quality of noisy speech and noisy images. The applications of SP are too numerous to list, but the depth and breadth of its impact are illustrated by several examples relevant to the NII. In Chapter 6 we return to more detailed consideration of a few of them.

Health Care Biomedical data acquisition and processing, including such sources as CAT scans, fetal heart monitors, EKGs, and high-resolution rendering from imaging devices. PET and SPECT illustrate associated inverse problems. Also, Picture Archival and Communications Systems (PACs) require all aspects of digital communications SP as well as prospective content-extraction algorithms.

Education The recent acquisition by NASA of the photos of comet fragments hitting Jupiter, and the dissemination of the data via the ethernet to scientists and grade students alike demonstrate the wide impact that image acquisition, storage, transmission, and display can have at all levels of education. Rapid access to new and old information through the NII can partially compensate for outdated textbooks and otherwise inadequate facilities.

Geographical Information Systems Huge databases are being generated for scientific, government, and commercial purposes that can be viewed as multidimensional maps of the earth and of the environment. Basic SP for communications and storage already makes these data more accessible, but enhanced SP tools will increase its utility for wider audiences.

Resource Exploration Oil and natural gas exploration commonly use DSP techniques to acquire and to process seismic signals. SP place a similar role in underwater and other geophysical signal analysis.

Modern Communication Networks Digital communications use DSP in modems, cellular telephones, commercial switching equipment, and teleconferencing.

Consumer Electronics Sophisticated DSP is used in compact disk players, audio processors, radios, computer graphics interfaces, and video games.

Automatic Navigation The older LORAN and more recent global positioning satellite (GPS) navigation systems are widely used for accurate navigation. Both systems use sophisticated SP algorithms to accurately estimate the user's position. Auto makers and publishers already offer prototype systems with on-line maps and travelers' directions.

Image Processing Image processing is used in many areas. For example, commercial silk screening now uses image processing algorithms to manipulate images, to separate colors, and to produce the masks used in the screening process. The fine arts have used SP analytic techniques to detect overpainting and to establish authorship and dating from radiographic images.

Entertainment SP is involved in many entertainment applications such as interactive video games and other multimedia applications incorporating sound, graphics, and imagery. Variations on some of these applications can support rather than detract from education. Multimedia access to information resources can be fun and can provide user-customized capabilities for motivation building, interactive measures of attentivity, and reward mechanisms for cooperative and successful behavior.

Business and Finance Videoconferencing enhances remote project planning, assessment, and collaboration. Multimedia methods are already used for remote processing by insurance companies, the real-estate industry, and direct marketing companies.

The previously described applications are already mature, although rapid and radical changes have occurred during the past decade.

4.5 A Signal Processing Example: Commercial Video Compression

Commercial video communication systems provide an excellent example of the importance of SP and the evolution of technology because such systems require a variety of specialized SP for video compression and decompression, audio compression and decompression, and the associated transport and control of compressed signals for consumer and private broadcast and video-on-demand applications. The industry viewpoint is short term (i.e., limited to effects of technology development on products and processes in the next five years). These applications tend to draw on recent developments in video and audio processing, modulation and networking technologies, and distributed computing.

Research ideas which are destined for widespread and economically important applications move through three stages:

1. The basic idea stage, dominated by research disclosure showing the possibility of substantial benefits, and significant interest and excitement in the research community

2. the refinement stage, in which the technology appears in a few commercial applications and is refined, establishing its practicality in gradual steps, and

3. widespread deployment stage, in which the technology appears in many applications and becomes taken for granted.

The basic concepts of much of the SP in emerging products are derived from research of the 1960's and 1970's (e.g., discrete cosine transform, motion compensation, sub-band coding, quadrature amplitude modulation, convolutional and Reed-Solomon codes). The technology refined from these ideas has reached a level of performance and implementability that has motivated significant investment in integrated solutions. This, in turn, is moving applications from niche markets to broad business and consumer markets, and is thus is a candidate contributor to the NII.

The research ideas mentioned previously are now moving from the second to the third stage. Many experts believe that in video compression and possibly audio compression, no ideas have moved from the first to the second stage during the past decade, nor are there any on the horizon. Two of the most widely followed developments in the last few years, fractals and wavelets, have not yet made it out of the first stage (at least as applied to compression). With this observation in mind, we now turn to the future.

Chapter 5

Signal Processing for the NII: Future

5.1 Introduction

Several important SP advances over the last few decades, such as modulation, error correction, and signal compression, help make the NII possible. However, given the maturity of many key ideas in these areas, research in the future should largely bifurcate between immediately practical commercial advances (the kind the Japanese relentlessly exploit) designed to move technology from refinement to widespread deployment, and the search for fundamentally new and different approaches designed to produce new initial ideas. In this chapter many topics of importance to future development are described with an emphasis on the development of novel ideas in conjunction with the user community of the NII and their movement into that community.

5.2 Compression

The current set of compression algorithms has proved successful for certain well specified tasks. Many standards are quite good and their existence has sped the development of the NII as well as public and commercial networks worldwide. Nonetheless, the need for further compression is as evident to any student who has tried to use Mosaic over a phone line as it is to an engineer trying to fit high definition television (HDTV) into the bandwidth currently available. New applications will push compression technology even further. Smart telephones coupled with personal computers and TV monitors accommodating both printed and video data will be common in homes and businesses within the next decade. Cellular telephone frequencies have become hopelessly congested in some areas. The requirements of efficient use of the NII and many related applications demand more efficient compression techniques for an increasing variety of applications.

Because compression comprises such a wide variety of techniques and applications, many promising avenues for future development are available. Several are reviewed here.

Off-Line Lossless Algorithms

As an historical artifact, lossless compression has largely ignored *off-line* algorithms (those that assume the input to be present all at once) and *super-linear* algorithms (those whose running time grows faster than input size.) But many applications impose no constraints on computation. Data in computer memory is there all at once; ever increasing memory sizes and decreasing costs permit communications to be buffered in huge blocks. Thus many applications now permit data to be compressed off-line. This is much closer to the ideal contemplated in Shannon information theory and to current practice in lossy compression.

Parallel and Distributed Implementations

We know of many highly successful, mature compression algorithms — successful in compression performance and versatility, mature in the senses that they are well understood, they have been tweaked to death, and almost nobody expects substantial improvements in compression. Examples include the LZ algorithms, in both lossy and lossless versions, as well as variants of transform coding and predictive coding. Some aspects of these algorithms can trivially be done in a parallel/distributed fashion — computing orthogonal transform coefficients, for example. Decomposing other tasks may not be straightforward, when all the computations are interdependent. Some wavelet and fractal transforms seem to be in this class; their enormous potential has been stymied by their computational loads. The NII provides a test bed for the distributed computation of such problems, and the NII would benefit from their solution.

Lossy Compression Algorithms

Incremental improvements in the existing standard schemes such as code excited linear predictive coding for speech, subband coding for audio, and transform coding for images are likely to continue. The use of the perceptual coding ideas described later plus methods of optimizing the free parameters such as the quantization tables in JPEG are likely to continue as they promise improved performance while retaining compatibility with existing standards.

Algorithms differing from the standards in fundamental ways are the subject of active research, but progress has been slow and sometimes controversial. To threaten a standard, an algorithm must promise either comparable performance at a much reduced price, much improved performance at a comparable price, or enhanced capabilities such as the incorporation of other SP to improve the utility of the compressed data. Active research areas include the following:

Wavelets

Many compression schemes include signal decomposition as part of the compression scheme, as the JPEG and other standards incorporate discrete cosine transforms prior to quantization. There has recently been intense research activity in a collection of methods for signal decomposition known collectively as *wavelets*. The theory and practice of wavelets unify several new and old methods, including quadrature mirror and perfect reconstruction subband filters which have proved useful in speech coding methods. A primary advantage of wavelet methods in comparison to the classical Fourier techniques is that they are naturally multiresolution and scalable in applications so that a single decompositions efficiently provides reconstructions at a variety of sizes and resolutions.

Wavelet decompositions provide an alternative to the sinusoids of Fourier analysis constructing a set of basis functions by scaling and shifting a single signal, or a few at most. By replacing the DCT of JPEG with a discrete wavelet transform (DWT), one has the potential for compression algorithms without the blocky artifacts of the standards at low bit rates and for achieving better performance in terms of tradeoffs of quality and bit rate at a user-selectable collection of bit rates — they are naturally progressive and multiresolution. Wavelets may also have advantages in terms of signal decompositions for subsequent signal analysis such as enhancement, edge detection, classification, and related statistical SP. Wavelet-based compression schemes also appear particularly well suited for cheap dedicated hardware implementation.

Research on wavelet-based compression algorithms is active at numerous academic and industrial research laboratories. Activity is focused on the wavelet decomposition itself, on scalar and vector quantization of the wavelet coefficients, and on low complexity software implementations and low cost hardware implementations. Variations on these schemes are likely to be serious contenders for future standards in image compression.

Quantization Techniques

Scalar quantization is well developed in terms of theory and application; the only flourishing research area is that of building cheaper and higher resolution quantizers for A/D conversion, and the theory of and practice of feedback quantization. Most activity in the theory and design of quantizers for lossy compression has focused on the design of quantization tables or bit-allocation algorithms for multiple scalar quantizers and on design algorithms for more general vector quantizers. In both cases the quantizers can be applied either directly to the signal or, more commonly, to a decomposed signal following a transform such as a wavelet transform or other preprocessing. When vector quantizers are used without signal decompositions the bit-rate versus image-quality performance can suffer, but extremely simple decompression algorithms involving only table lookup and no computation can be used. This provides an attractive alternative in applications such as decompressing video on a personal computer where high quality is not needed, but fast and simple software decompression is required.

Vector quantizer design algorithms have concentrated on quantizers with the geometrical structures called lattices and trellises and on quantizers designed using statistical clustering methods. Lattice vector quantizers and trellis coded quantizers have proved useful for medium rate speech compression and they have shown promise for low bit-rate image compression when used in cascade with wavelet decompositions, providing good quality images at under .25 bits per pixel. The clustering methods have the intriguing side benefit of resembling statistical classification algorithms and hence suggest the possibility of combining compression with other SP such as classification and segmentation. A variety of clustering algorithms have been reported in the literature, including ideas from neural nets, self-organizing feature maps, simulated annealing, and cellular automata as well as many algorithms long used in statistics and taxonomy. Although active research continues in these areas, there has rarely been clear evidence of any general superiority of one approach and comparable performance is achievable in a variety of ways. Most significant advances have come from either matching a quantization technique to a particular signal decomposition such as the DCT or DWT, or from imposing a low complexity structure onto the quantizer and then doing a constrained clustering design.

A form of vector quantizer that has received much attention in the popular press and the computer magazines is the iterated-function system or fractal code. These codes replace clustering or lattices by an algorithm which finds blocks within an image which can be used to approximate other blocks within the image by simple operations. Decompression is simple because it can be considered as a vector quantizer operating on raw pixels. Results, however, have been at best competitive with the JPEG standard in performance, while requiring higher complexity. Although fractals have proved extremely useful for artificial image generation, they have yet to have a major impact on the compression of natural images.

While neural networks have not shown marked advantages in designing vector quantizers, they hold promise as a means of implementing the quantizers in hardware. Neural nets have also been touted for the “learning” implicit in adaptive compressors, but evidence is scant.

Universal and Adaptive Compression

An important direction for research is adaptive or universal compression, compression techniques that can change to accommodate differing behavior within a signal. Recent theoretical advances need to be translated into practical coding schemes, requiring substantial efforts. Desirable goals include fully adaptive transforms and quantization schemes.

Multiresolution and Scalable Compression

Multiresolution and scalable coding schemes will be central in multiuser environments, especially when those users have widely differing degrees of access to the NII and varying computing and display facilities. They are also likely to play an important role in algorithms for browsing since low resolution images provide an effective means of browsing image databases.

Multisignal Compression

Many signals have multiple components: color images have red, green, and blue components; multispectral LANDSAT images have differing images for each segment of the electromagnetic spectrum; medical images are often taken as studies incorporating many images of a common organ, possibly using differing modalities; combined voice/video coding for videophones and multichannel audio joint coding involve multiple signals simultaneously. Better techniques must be developed for taking advantage of this structure then do existing codes which separately code each component without regard to its context. Further examples include combined voice/video coding for videophones and multichannel audio joint coding (the new HDTV audio standard has 5 channels).

Hardware

Practical use of compression techniques requires developing special purpose hardware that reduces the weight/power requirements on the compression module while at the same time reducing per-unit cost of the module. The commercial chips recently developed for implementing various compression techniques at video rates include two-dimensional DCT chips with and without quantization and Huffman coder, DPCM chips with Huffman coder, two-dimensional Hadamard transform with Lempel-Ziv chips for lossless compression, and Wavelet transform processor chips. In addition to these, there are a number of vector-quantization chips (using tree structured codes with pipelining) and neural-network chips which operate at rates well below the data rate for real time TV.

Accommodating high definition TV on the NII network requires the development of high speed chips and special purpose processors for all aspects of compression, including signal decomposition, clustering, and quantization.

5.3 Perceptual Coding

Compression schemes often operate on signal values like the amplitude of speech at a specific instant (sample) or the intensity of an image at a specific location (pixel) without regard to the way that the final reproduced signal will be heard or seen by a human user. This is appropriate for some data such as measurements or text, but it fails to take advantage of potentially useful information when reconstructing a signal intended for subjective perception by humans. If, for example, greater compression can be achieved at the cost only of loss imperceptible by the human ear or eye, then a lossy system can appear to have as high a performance as a lossless system with far inferior compression. Compression methods taking advantage of the nature of these phenomena are referred to collectively as *perceptual coding*, and seminal work during the past decade promises significant improvements in compression. Perceptual coding can be accomplished by a variety of means, but it usually involves using models of human perception, such as a human auditory system or human visual system model. These models can be quite complex and their incorporation into compression algorithms quite involved, often involving cooperative work among psychologists, computer scientists, and engineers. The potential gains have been estimated at 10–50% improvements in efficiency of compression with no perceptual distortion. One approach is to transform the raw data using the perceptual model into features deemed important for perception. It is these features that are then explicitly compressed and used to reconstruct the signal. Another approach is to incorporate the perceptual knowledge into the measures of distortion and fidelity used to design the codes. Regardless of the specific method, sensible incorporation of quantitative aspects of human perception is likely to provide substantial improvements in compression performance for speech, audio, images, and video with a modest increase in cost or complexity.

5.4 Quality and Utility Evaluation

Closely related to compression and perceptual coding is the issue of measuring quality or utility in lossy compressed signals (or in any signals which have been altered by SP) in a meaningful way. The shortcomings of common tractable measures (squared error and signal-to-noise ratio) have long been known, but thorough, statistically reliable testing of subjective quality using human subjects is expensive and not always conclusive. There do not exist generally acceptable or standard methods for demonstrating that a particular SP algorithm indeed provides an improvement in either perceived quality or utility in an application. A variety of methods have arisen in different fields for quantifying subjective quality in speech in images, diagnostic accuracy in medical images, and related issues. Such methods must be developed cooperatively by users and algorithm developers, often with the assistance of statisticians to provide sound experimental design and statistical analysis to replace the frequently anecdotal evidence given in support of algorithms. Since formal validation of algorithms can be expensive, involving extensive tests with professional users, it is important to develop easily computable, quantitative quality metrics that can be used to compare candidate algorithms and to improve their design. Such metrics are likely to depend strongly on the accuracy of models of human perception.

The availability of official or *de facto* standards for evaluating the quality and utility of algorithms that change data would greatly enhance the ability of the technical and user community to emphasize genuinely promising techniques and to limit the influence of unsubstantiated fads.

5.5 Communications

Important questions in signal design must be resolved to improve performance in the wireless portions of the NII. They include code division multiple access (CDMA) versus frequency division multiplexing (FDM), or time division multiplexing (TDM), power control, adaptive equalization and antenna processing. The use of array processing and beam forming for mobile communication is one example of success. Wireless connections to the NII could provide significantly improved access from portable computers and remote sites.

The broadcast channel for digital delivery is another topic for further investigation, given the importance of both digital television and audio broadcast.

5.6 Joint Source and Channel Coding

Shannon identified the major components of SP for communications and storage as source coding, channel coding, and secrecy. He proved, moreover, that in the ideal these could be solved independently and then combined for optimal performance. Researchers sometimes forget that this independence holds only for restricted models of source and channel, and at the price of infinite delay and infinite computing power. In practical situations, joint coding can yield significantly better performance. Improvements have been demonstrated in modulation, but remain to be seen in the combination of compression and secrecy.

Delivering real-time multimedia services to a mix of users over possibly congested networks requires innovative solutions that have only begun to be investigated. Worst case scenarios (using bandwidth reservation) or "best effort" solutions (which break down under serious congestion) must be improved if efficient multimedia communication is to fulfill its promise.

A possible solution uses multiresolution source coding (quantization and compression) methods combined with transmission schemes providing different grades of services (e.g., priorities). Such approaches have yet to be studied in detail and optimized. As an example, video services are currently available over the Internet on an experimental basis. They use a multicast backbone (MBONE) to broadcast video and sound to groups of users. There is no bandwidth reservation or resource management by the network. The network video tool (*nv*) from Xerox PARC is based on an intraframe coder and the INRIA videoconferencing system (*ivs*) is based on H.261. Neither uses joint design of coding and transmission, and both are sensitive to network conditions. Substantial work on joint design of real-time audio/video coders for packet services is needed in order to achieve good quality and efficient use of networking resources.

5.7 Image Processing

Here are several important SP issues primarily or uniquely concerned with image processing:

Still Images from Video There is a great demand for creating high quality still images from a sequence of lower quality video frames. This task will require research in the area of motion detection/estimation in addition to noise and blur removal and other generic image processing algorithms.

Image Understanding Algorithms that accurately simulate the SP steps employed by the human brain for image perception can help automate many of the currently costly (because of the need for intelligent operator intervention) image processing tasks. Smart algorithms might determine image orientation, perform contrast enhancement prior or subsequent to image capture so that the picture taker is assured that "what you see is what you get" (WYSIWYG), locate and remove red-eyes due to flash exposure in a photograph, etc. Needless to say, any advance in image understanding research will directly benefit other areas of image processing such as image compression (object-oriented compression, visual-based compression models), image restoration, and image enhancement (object-selective enhancement).

Compression of Compound Images Many compression standards such as JPEG and MPEG are geared towards color or grayscale images while the CCITT fax standards and the emerging JBIG standard target text and graphics. We need a universal algorithm to combine the benefits of both approaches to successfully compress compound images. The need for such an algorithm in a multimedia environment can not be overemphasized.

Image Hierarchies Multi-use environments, where an image is accessed at various resolutions, dictate the need for the construction of variable-resolution pyramidal image hierarchies. For example, the Kodak PhotoCD system uses a simple method of constructing an image pyramid (with resolution levels from 6152 × 4096 down to 128 × 96) by compressing the residual information at each level of the pyramid. In the same vein, the use of progressive transmission techniques, which reduce the effective transmission or retrieval time from image databases, necessitates the need for fixed-resolution, varying-detail image hierarchies. Improved techniques are needed to minimize the overall entropy of the image hierarchy (and thus its bit rate) while maximizing perceptual quality at each stage. These methods have been considered in Chapter 5 and elsewhere, but they are particularly important in image processing.

Color Management System As images get transmitted across networks there is an ever increasing need for a color management system that would enable users to reproduce a color image with its true colors.

5.8 Imagery Data Fusion

Areas of SP that must be developed to make a reality of automatic real time image data fusion include:

- Edge/Motion Detection Algorithms
- Multi-Resolution Processing Algorithms
- Segmentation and Classification Algorithms
- Shape and Object Recognition Algorithms

5.9 Intelligent Information Agents

Information agents are programs that automatically perform customized information processing actions to deal with "information overload" problems. Examples of agents are programs that automatically sift through mail, filter bulletin board postings, search out and organize documents in large databases as requested by the user. At their lowest levels of operation, information agents must correlate data typically in the form of language text. In this framework, documents such as email messages, reports, notices, and so on are signals whose representations are abstracted or compressed in some way. A typical approach to text representation for correlation is the so-called "vector space" model in which word frequencies are used to distill the content of a document in a crude way.

The representation of a document in this model, which is currently state of the art in information retrieval, is a very long, sparse vector whose coordinates represent dictionary words and whose coordinate values are the relative frequencies of those words. Given the size of the grammatically nontrivial vocabularies that arise in most applications (even when specialized), this leads to vectors with many thousands of coordinates. These vectors must be compared, correlated, clustered and otherwise statistically processed in order to organize and retrieve appropriate information and take some subsequent action. Such operations are standard in SP and many techniques, algorithms and models can be transferred between fields. Statistical SP is already playing a role in some approaches wherein principal component analysis of a document database is used to eliminate "noise" and establish synonymy automatically. This technique is known as Latent Semantic Indexing (patent pending by BellCore).

From the SP operation on a signal family, essentially taking a Karhunen-Loeve decomposition. It is evident however that research on information agents should be closely coupled with SP theory, modeling and algorithmics to leverage the experience and success of methods already known there.

5.10 The Human Interface

SP has not yet delivered as much as it should in supporting ease of human interaction with the NII. To be truly democratic, the NII will need to be accessible without extensive training and special skills. Some key areas of SP research, such as speech recognition, could make an enormous impact but have been frustratingly slow in delivering useful results. Apple's *Newton* showed the immaturity of handwriting recognition. There will be an enormous need for translating different source representations to be usable on a wide variety of output devices. For example, how do you get the useful information from a large mixed media file over a low data rate link to a limited capability palmtop terminal? Browsing signal databases might be the answer. Given the average American's frustration with programming a VCR, one might expect enormous benefit in research to help plug our citizens into the network we are creating.

In addition to the specific topics considered in this section, many of the subsequent sections can be considered as topics relating to improving the human interface.

"Reading" Information in the ¡NII

Getting information over the Internet and "reading" it may often be less efficient, less smooth, and more frustrating than reading the information out of a book. When we read a book in the normal sequential manner, the technology of bookbinding imposes itself only to slow us down when we turn the page and it slips. We lose 3 seconds in the middle of a sentence — a minor inconvenience every 5 to 10 minutes.

When we "read" hypertext information (with images) using Mosaic, we are interrupted more frequently, and for longer periods of time. Clicking on a hotword with a mouse takes roughly the same amount of coordination as turning the page of a book. However, waiting for the download of information may take 10-20 seconds.

Interaction with the informational material in the NII needs to be tighter than it is currently. SP techniques can contribute to timeliness of responses in "NII reading," particularly where images are concerned. Interaction can be much richer than "turn the page" or "download the image with this filename." NII technology and practice does not support the analog of skimming an encyclopedia article. The faster one reads, the higher the percentage of the user's time is taken up waiting for downloads. Thus SP researchers should be involved in the process of developing new interactive techniques for inspecting images at a distance and for browsing multimedia databases.

As long as we deal with pure text, SP is not so relevant. CS and AI systems for smart lookahead, caching, loading multiple pages at once, etc., can handle the problem. Things are different when we deal with images, video, audio, or graphics other than pure monofont text.

Here are some specific techniques (existing or potential) which may enable effective Internet reading of the sorts discussed above:

Knowledge-based Compression For human-face images we may use a special sort of knowledge-based compression: before transmitting an actual scanned image, compute and transmit a "pen-and-ink sketch" of the face, and transmit that first. Follow this with the full scanned image for progressive transmission.

Content-based Progressive Transmission Using knowledge about what the important objects and visual features are in each image to be transmitted, an intelligent progressive transmission system can transform and order the image information to deliver the most information-laden bits first.

Interaction-based Progressive Transmission When knowledge is not available to the system to direct a progressive transmission, information such as mouse position, etc., coming from the user in real time can help direct the refinement of the image into details.

1. The user consciously directs the transmission of details.

2. The user subconsciously directs the transmission of details. This might involve corneal-reflection eye tracking or something similar.

Low-resolution Look-ahead All images potentially available to a hypertext document via a single link can be prioritized with a heuristic value that models both the likelihood of being followed and the importance of that information being quickly available. A series of requests would be issued by the client software to retrieve the top K of these N images.

User-friendly Multimedia

As the NII grows, enormous quantities of multimedia (MM) information becomes available to an expanding community. Today's MM information handling is rather primitive, involving many manual operations for creation and assimilation. There are few tools for combining MM objects and data into forms that are both meaningful and pleasing for the user. With the exponential growth of MM providers, all vying for the attention of users, such capabilities will be essential.

Examples of research and development directions that will promote such capabilities follow.

Automated MM Specification We need an automated way to identify and describe the content of audio, video, text, graphics, numerical data, software, etc. This should involve as little human interaction as possible. Algorithms for MM understanding should dissect the MM objects into constituent parts, identify them and provide labels. Required technologies include audio speech recognition, video scene analysis, text concept identification, software capability detection, automatic closed captioning, etc.

It is important to note that there can be no really effective substitute for careful content description by knowledgeable humans using standardized description languages. But this is relatively expensive and is likely to lag behind automatic indexing processes, which thus have a role to play simply in providing a fallback means of classification or image retrieval. Furthermore, automatic techniques serve as tools that boost the productivity of human content analysts.

Automated MM database Creation Given descriptions of MM objects, create databases and catalogs with index information. Automatically search for relationships between MM objects and include these in catalog. Distribute catalog index information. Check for obsolescence, and delete old objects.

Automated MM Reviewing Automatically determine the quality, technical level, timeliness and relevance of each MM object. Authenticity should also be checked at this stage. Produce metrics for novelty, utility, redundancy, boredom, etc. tailored to the individual user. Feedback to author/creators of MM object as a straw poll of usefulness.

Automated Program Assembly Database search should discover nonredundant, high quality MM objects of a technical level matched appropriately to the user. These should be combined, using "hyperglue," into a meaningful, enjoyable and exciting program.

Machine Assisted or Automatic Authoring/Creation of MM objects Database search for MM objects useful for current creation. Interaction with quality checker to maintain utility and nonredundancy. Check that all technical levels are covered.

Creator/User Control of Security and Privacy Tools for detecting counterfeiting and impersonation. Key control mechanisms. Public and private sections of MM objects.

MM Gleaner Tools for audio/video speedup or slowdown. Audio gap removal.

Congestion Sensitive Service If system is congested, avoid large downloads. Tailor service to congestion.

Asymmetric Transmission Services Downstream via CATV or satellite, upstream by modem. Significant computer power available on one side (sophisticated software or hardware-based compression), limited computer power on the other (need fast software decompression).

Better MM Scripting Languages Enables remote downloading of MM processing software. Need set of standard primitives, e.g., playMpeg to decompress video compressed using the MPEG standard.

"Smart" Signal Acquisition Speech enhancement for noise removal in mobile telephones prior to compression. Voice activity detection (pattern recognition task) for variable rate and packet voice systems.

Random Access Retrieval from Compressed Databases This may require new compression algorithms to allow easy decompression on the fly from any point in a large compressed data stream representing, for example, combined audio and video. It may be necessary to browse through such databases to locate specific items, e.g., a particular spoken phrase or particular person's face. Word spotting/locating (or "gisting") in compressed speech file. (An audio analog of the Unix "grep" function.)

Sound synthesis for music and other special effect sounds Computer based music and sound composition; video games and other CDROM based applications — a major commercial area of application, currently where the big bucks are in multimedia for U.S. industry.

Image and video transformations and synthesis Interactive design of faces and full body models of people. Signal and image processing for virtual reality and movie animation. Wire-mesh human head models and dynamic models of face and lip motion. Speech synthesis coordinated with lip motion dynamics.

Speech recognition and synthesis Much of speech recognition and synthesis is SP, and these remain active research areas. Existing recognition algorithms remain constrained by artificial grammars or limited vocabularies and usually require long training for individual speakers.

5.11 Enhanced User Functionality with Images

In 1982 the French writer Henri Hudrisier described his vision of public centers where collections of culturally important images would be available to researchers, writers, and artists. These centers, called "iconotheques," were to provide not only the images themselves, but facilities for working with them, including workspaces for visually comparing images, tools for creating documentary presentations using the images, and an environment for discussing the images with other people. Hudrisier believes that the proliferation of images in 20th century western culture has brought a revolution in the "mentality" of the people, and that images are continuing to grow in importance as shapers of human thought. As people grow in their fluency with images — creating presentations and witnessing presentations, they attain a new kind of literacy with exciting possibilities for human expression and understanding — a welcome relief from some of the more unfortunate aspects of the information society (e.g., unemployment as a result of automation).

The NII has the potential to make any citizen's PC screen become both a personal iconotheque and an entryway to a vast, global iconotheque. Standing in the way are three kinds of obstacles: current access limitations in terms of bandwidth and equipment, lack of appropriate software tools and interfaces, and lack of human readiness to exploit iconotheques. The SP research community has potential contributions to make towards the solution of each of these three problems. Bandwidth limitations are addressed by work on compression, progressive transmission, etc. Software tools include audio, image, and video processing and analysis, database tools and editing systems. These obviously are intimately tied to digital representations of signal information. In terms of human readiness, the SP community is poised to contribute to an educational effort that will not only acquaint citizens with concepts of digital representation, resolution, bandwidth, etc., but with many of the possibilities for interesting transformations and reformatting of signals in support of artistic expression, special effects, effective documentary communications etc.

5.12 Telecommuting

The key to telecommuting is capturing the essence of interpersonal communication present in a traditional office setting that is necessary for a job and replicating it via communications and SP technology. The more jobs that can be adequately handled, the more we can relieve urban congestion and related ills. The SP work relating to human-machine interaction mentioned earlier is directly relevant. Also intriguing is cross disciplinary research (e.g., sociologist/psychologist/electrical engineering) targeted at telecommuting. Xerox PARC and Bellcore have conducted some such research.

5.13 Signal Processing for Mining Information

While the NII allows access to orders of magnitude more information than previously possible, the question remains how to find the information of particular interest to a user.

Processing documents for filing, classification and storage requires advanced segmentation, optical character recognition and image compression techniques. Clustering methods for organizing data, including images, are relevant. Image processing, computer vision and image understanding techniques will be used for intelligent indexing and classification, for example of large distributed image or video databases. Intelligent agents searching for relevant information will require advanced SP techniques for finding attributes or extracting features when sifting through multimedia data. Mining information overlaps the topics of automatically indexing multimedia objects for database creation and information agents. It could be included under the "Human Interace" heading, but it is sufficiently important on its own to be highlighted separately.

5.14 Signal Processing and Virtual Reality

Virtual reality (VR) is a combination of technologies whose goal is to integrate real and synthetic sensory signals and present them to a user in such a way as to evoke the sensation of being immersed in a virtual environment. Early applications of virtual reality are showing promise in design, manufacturing, medicine, data navigation, entertainment and training. While the promise and promotion of this technology is high, it is still largely a laboratory concept and much work remains to be done. Nonetheless, it is easy to imagine an increasing amount of traffic on the NII being devoted to data in support of virtual environments, with distributed computational and sensory resources contributing to a single virtual environment.

The SP problems in virtual reality are novel and relatively unexplored. Solutions to these problem are critical to the success of VR. A simple but obvious example is the processing of stereo sound so that sources can be arbitrarily relocalized in a synthetic audio field. A similar question can be asked for stereoscopic vision — how can three dimensional scenes be represented and compressed effectively so that the end user can efficiently simulate an arbitrary view? Are there advantages to combining stereo audio and video signals? What are the signal representations most effective for overlaying synthetic images on real images accurately and efficiently. What are the properties of tactile senses (texture, etc.) and how can they be effectively measured, compressed and reconstructed?

5.15 Standards

At the current level of maturity of enabling technologies, standardization becomes increasingly important owing both to customer interest in interoperability (for convenience and to promote vendor competition) and the price reduction made possible by sharing common components over multiple vendors and applications. Standardization may be either de-facto standardization around a proprietary technology (e.g., General Instrument's *DigiCipher*), consortium standardization (e.g., ATM Forum), or formal standardization (e.g., the International Standards Organization's MPEG).

Several of the important new standards are fairly complex and imply significant software content. It is not axiomatic that U. S. companies are at a disadvantage in competing when such standards emerge. It may be worthwhile to consider a new generic strategy towards standardization in which we collectively target research with the goal of supporting particular standards and accumulating in-depth knowledge of the technology and implementation issues. The research would be conducted across academia, government, and industry, and could be focused by the NSF as described above under the topic "Use the NII to foster the NII." Funding activity would assist such focus. The goal would be to foster truly excellent standards and permit U.S. implementors and users to lead in creating and exploiting applications.

5.16 Education and Libraries

Education and libraries can benefit both from SP progress on human-machine interaction and on the problem of multi-media database browsing. It is unlikely that all schools and libraries will have access to the same quality of computing and audio-visual equipment, yet we would like to provide access to the essential elements of our archives to all, creating the need for translation of complex multi- media presentations to simpler formats that can be displayed and manipulated on terminals of lesser capability.

Another important educational issue relevant to the NII is that of educating the public about the NII and its enabling technology, including SP. The change to an increasingly digital world is fundamental and needs to be understood by every citizen. Introducing the NII without educating its users is a formula for disaster, because digital representations can be easily misinterpreted.

For example, how does one judge the veracity of a picture in a newspaper, on TV, or in an electronic journal? What if a prankster (or worse) perpetrated fraudulent news of an attack, or an assassination, by putting out a doctored image along with a textual news break. People to whom "seeing is believing" might well panic as they did during the Orson Wells broadcast of H.G. Wells' *War of the Worlds.* A related question asks "Is it right to admit a digital image as evidence in a court of law? If so, then under what circumstances?"

What is resolution? How does it relate to the "reality" of an image or sound recording? Should high-resolution images be treated differently as legal evidence than low-resolution images? How does resolution in images and audio records relate to resolution and granularity in other representations of the world (e.g., simulation models)?

The following concepts related to resolution would ideally be understood at least qualitatively by all citizens of the 21st century:

1. Digital sampling, the Nyquist limit, aliasing,
2. Quantization and bits/sample, bits/pixel,
3. Frequency analysis (e.g., Fourier) and the difference between the frequency domain and the space or time domain, and
4. Bandwidth.

5.17 Image Analysis

Many applications of the NII incorporate some form of automatic image analysis. These techniques come from a variety of sources, including traditional and modern statistical classification and regression theory and practice and the mathematical modeling of objects and their automatic recognition by statistical or morphological methods. Examples include digitizing scanned documents via character recognition, searching databases for specific features or context, and automated "second opinions" in medical screening and diagnoses.

Parametric approaches involve fitting mathematical models to the data and using the extracted parameters to make inferences based on the theoretical properties of the models. Nonparametric methods generally involve few or no *a priori* assumptions on the structure of the data; they learn the structure needed for inference from training sets of typical data. A variety of methods have shown promise for modeling in image analysis, including random fields, stochastic morphology, and stochastic partial differential equations. Research on such model theory is not likely to have an immediate impact on SP for the NII, but over the long term such methods enrich the general development of SP.

Nonparametric approaches have recently been successfully applied to raw image data or extracted features in order to classify specific features in imagery such as anomalous tissue in medical images or wheat fields in aerial photographs. Such methods include nearest neighbor classification, statistical clustering, classification and regression trees, linear discriminant analysis along with the modern extensions to penalized and flexible discriminant analysis, empirical Bayes classifiers, and artificial neural networks. These methods can assist in image segmentation, analysis, and enhancement and can perform "compression" by extracting only those portions of images important to the user.

Most of these methods are the subjects of active research in a variety of fields such as Electrical Engineering, Statistics, Computer Science, and Medical Physics. Although a useful stable of basic tools exists for simple problems, much remains to be done for most complex tasks found in the real world.

Chapter 6

Application Areas

6.1 Introduction

In the following sections several specific applications areas are explored in further detail and many of the previously considered themes are developed in the context of particular applications.

6.2 Geographical Information Systems

Background

The field of geographical information systems (GIS) historically has had little overlap with traditional signal processing. Such systems deal with geographic data, which can be in one of two forms: raster data or vector data. Raster data are of the form treated by traditional SP; e.g., pixel intensities in a scanned sampled image. The diagnostic of "vector" is the presence of coordinate pairs, describing point, line, or area features (lines and areas as sequences of vertices assumed connected by straight lines). Features are associated with (often large) numbers of attributes, e.g., text representing a highway number or lake name. Good examples are digital datasets derived from U.S. Geological Survey (USGS) quad maps, or digital representations of street networks. Sometimes raster datasets (earth images) are created directly, but more often they have been digitized or scanned from paper maps. If they have been scanned, raster/vector conversion and generalization have likely removed any trace of the raster. The equivalent of the compression problem for these data sets is generalization, i.e., the weeding or pruning of vertices, or more advanced techniques. Quadtrees are used along with other types of trees as indices for access to features.

Characteristics of Geographic Data Sets

Geographic data have explicit ties to locations on the surface of the Earth. Thus they form a subset of spatial data, which have any two dimensional reference frame, but the two terms are often used interchangeably, and "geospatial" is recently popular.

The NII has much to offer the field of geographic data. Historically, geographic information has been cumbersome to handle (maps, hard copy images, atlases) and is not addressed well by conventional catalog systems. Thus map libraries are a class of "special libraries," physically and intellectually separate from mainstream libraries. Conventional searching systems typically require physical access to materials, which are stored in cumbersome map cabinets. Difficulty of access is a major factor in explaining the generally low level of utilization of geographic data, and perhaps also the low level of geographic knowledge and awareness in U.S. society.

Efforts began perhaps three years ago to define a National Spatial Data Infrastructure (NSDI), within the umbrella of NII. A National Research Council (NRC) report *Towards a Coordinated Spatial Data Infrastructure* appeared in 1993, and the Administration declared a commitment to NSDI in the report of the Gore Commission, *Reinventing Government.* An Executive Order in April 1994 mandated the first concrete steps in building NSDI.

Currently there are three major efforts under way under NSDI. The first concerns the establishment of a metadata standard, to address some of the shortcomings of conventional cataloging of geographic materials within the context of digital information exchange. The second is a commitment to build a national geographic data clearinghouse, a "virtual catalog" that will help users locate suitable data. The third concerns the identification of key data sets that will form the "framework" of NSDI, allowing producers of geographic data to tie their information to the surface of the Earth at appropriate levels of positional accuracy. The framework includes the geodetic control network, and digital representations of hydrographic, topographic, and transportation features. Many other research issues can be identified that bear on the prospects for achievement of NSDI's goals. Some of the more technical aspects that potentially relate to SP are discussed next.

Research Issues

Geographic data come in two forms — images, and collections of digitized features. The latter are almost exclusively represented in vector form, and may have numerous detailed attributes. For example, the description of a polygonal object from a soil map may extend to thousands of bytes. Many vector data sets are "planar enforced," meaning that every point in the plane lies in exactly one polygonal object. This has led to a preference for data structures in which the common boundary between two polygons is the primitive object, rather than the polygons themselves. Geographic data is almost exclusively two-dimensional and static. Information on terrain elevation may be described as "3D," but because of the constraint that the Earth's surface has a unique elevation at virtually every point, it is traditionally treated as 2D, with elevation as a single-valued function of location.

Vector data sets include digital representations of topographic maps ("digital line graphs"); thematic maps of soils, geology, land cover, or vegetation; and digital representations of transportation or hydrologic networks. The discrete features in such data sets are represented as collections of point, line, or area digital objects, and tagged with attributes. The use of coordinates is a common diagnostic feature of vector data sets.

Raster data sets include remotely sensed images, regularly spaced samples of terrain elevation ("digital elevation models"), and scanned digitizations of maps. In most data models each cell is restricted to a single value, so multivariate cases must be handled as collections of "layers."

Tight Integration of Images and Features

The separation of raster and vector data mirrors the traditional separation of maps and images, and has been preserved into the digital era. The distinction is also preserved by software, and to date there are no geographic information systems that implement processes combining both raster and vector data. Many packages implement both classes of data structure, but processes are either raster or vector, requiring the user to be aware of the distinction.

With raster data it is impossible to find features, except by the implementation of on-the-fly pattern recognition techniques, and few geographic features are recognizable in Earth images (consider the query "find the Mississippi River"). With vector data, it is impossible to obtain more than the digital representation of a feature. The Mississippi River may be extracted as a polyline with attributes, but it is not possible to learn any more about the Mississippi than its footprint, and the attributes given it by the data producer. By contrast, image data almost always holds the possibility of more knowledge through interpretation or pattern recognition.

This high-level separation of image and feature-based data seems to be a significant and unnecessary impediment to applications of geographic information. A user should be able to search for "Mississippi River" and extract both an image and an interpreted vector representation. Thus far, methods to couple image and feature more tightly have not been developed. We need a new concept of integration of raster and vector data at the feature level. This reinforces the general goal stated earlier of developing imagery fusion methods for distinct types of information.

Support for Browse

Geographic data tend to be voluminous, and one reason for historical emphasis on vector data is clearly the potential for greater information content in more highly interpreted data. Although much vector geographic data originates with images, particularly air photos, a complex process of human pattern recognition and generalization occurs during the conversion to vectors. At this time, we are far from having the ability to automate either process. Automated generalization has proven to be a particularly thorny problem, and vector data sets of the same area are still produced largely independently at each published scale. Better methods of automated generalization and pattern recognition for geographic features and their attributes could reduce the costs of producing map series and vector data sets substantially.

A full Landsat scene is approximately 300 mB in uncompressed form, while a single quadrangle of feature data contains between 10 and 100 mB. Standard lossy and lossless compression techniques achieve substantial reductions for image data, but fail to take advantage of the inherent economy of feature representation. The tight integration of raster and vector described above may achieve much more effective information transfer per byte, and allow the development of efficient forms of browsing over the net. Simple expansion of bandwidth is not likely to produce the same results, at least in the short to medium term.

6.3 Planetary Science

An application in planetary science provides an example of the importance of SP in scientific exploration and analysis. The acquisition and study of data from the Galileo mission is dominated by issues of access to, and analysis of, imaging data. The issues have strong implications and lessons for planning the NII.

NASA's Galileo Mission to Jupiter

Spacecraft and telescopic imagery require archiving, access, and image processing that help define the requirements of SP for the NII. The scientific use of data from the Galileo mission is instructive. The Galileo spacecraft is now en route to Jupiter after innovative imaging of the Earth/Moon system and of two asteroids. Into the early 1980s, the nominal Galileo imaging data product was to be photographic hardcopy of roughly 100,000 images, with images acquired by a CCD camera and transmitted to Earth in digital form. Eventually the Galileo data plan changed to one based on digital data distribution. The data are well suited to distribution on digital networks like the NII, and image processing is usefully carried out on the original data sets by scientific users. Ironically, in addition to the many delays in the project, the failure of the spacecraft's high-gain antenna to open properly means that the data rate from Jovian orbit will be about 40 bits/sec. At that rate a full 800×800 pixel 8-bit image would require nearly 40 hours to return to Earth.

Thus the data conduit from Galileo is more like a snail trail than an information superhighway. The planned 100,000 images cannot be returned to Earth. Even if only 1000 are to be sent, we need to develop strategies for data compression and pre-processing. Thus, many of the same issues most be addressed as for the NII.

In all of the efforts to obtain optimal data compression schemes, one limitation has yet to be surmounted: No scheme seems able to compress image data by more than a factor of 10 and still permit quantitative scientific interpretation. This limit is a fundamental challenge to technical preparations for the NII, to the extent that it will be used for archiving and distributing scientific imaging data.

6.4 Health Care

SP research and development which will have a significant impact health care applications on the NII in four categories: technology, algorithms, databases, and interactive teaching.

Technology

Communication Technology

ATM (asynchronous transfer mode) is a recent advance in communication technology using circuit switching. This technology allows high speed transmission of large data files including audio, images, and video in both LAN (local area network) and wide area networks (WAN). Although the ATM standardization is still being finalized, it is anticipated that it will move forward very rapidly. With the inexpensive ATM connection available at each node in LAN and WAN, this will allow high speed communication among the users through the NII. During the past few years, LAN R&D is advancing much faster than that of WAN. Within universities, hospitals, and private industry, the speed and cost of using LAN for information transfer are acceptable for most data types other than images and video. However, once the information leaves the central location, the high speed WAN becomes inaccessible because of its excessive cost.

Several recommendations not directly involved with SP, but important for improving the general environment for the NII can be made:

1. To speed up the standardization process through the ATM Forum.
2. To get access into the cross-country fiber optic cables from cable carriers so that the cost of the wide area network becomes affordable.
3. To research and develop “ramps” between the local area network and the wide area network.

Image Compression and Coding

Existing standards such as JPEG and MPEG1 do not provide sufficient quality for many medical images at compression ratios of more than 10:1. It is important to develop improved lossy compression techniques in software and hardware which preserve quality while obtaining higher compression. This involves supporting the development of new compression algorithms and technologies, the establishment of large databases pertinent to each type of image modality in the health care field which could be used as a reference standard to evaluate different types of image coding techniques, and the investigation of 3-D and 4-D image compression methods that arise in modern biomedical imaging applications, such as radiology, surgery, and neurology. Multidimensional image compression is a variation on on the multiple source compression described in Chapter 5.

Storage Arrays

One major accomplishment in the field of SP is disk array technology. Three years ago, a five gigabyte disk array with a 5 MB/sec sustained I/O transfer rate cost over $60,000 and the array management software was difficult to use. Within half a year, a storage array with half the price and double the capacity will be commercially available. In addition, such a storage array will be effectively a peripheral device without any change of the data format. The large capacity and low cost storage array will be a very important ingredient in the national information infrastructure. It is important that research continue to improve the I/O speed, to increase the capacity, and to lower cost of storage arrays.

Displays

During the past several years, advances in display technology have been in the 1,000 line color, black-and-white, and the 2,000 line black-and-white display monitors. The brightness of the 1K monitor can reach 200 lambert while the 2K black-and-white monitor can reach 120 lambert. The frame rate can go up to 72 cycles per second. The 2K display image quality is excellent for routine clinical diagnosis in radiology and in the publication industry. However, for about 5-10% of the images in diagnostic radiology, including mammography, some bone diseases and lung diseases, higher resolution, 4K for example, is needed. There are currently no R&D efforts to investigate such high performance display systems. In addition, we are almost at the limit of CRT technology.

Algorithm Development

Three areas of research in two-dimensional SP algorithms would directly benefit the health care field:

Segmentation

Although many excellent segmentation packages are available for solid objects, many of these techniques become useless when the boundary between the object and the background is fuzzy. For example, segmentation of the brain in MRI and the lungs in digital radiography are still being done interactively. Currently, there is no systematic way of segmenting objects with fuzzy boundaries.

Extraction of Image Content

Image content of an object can be either the texture or some quantitative measurements within the object in the image. During the past few years, some texture analysis algorithms have been developed for artificial objects which perform reasonably well. However, when applied to real-life patterns, these textured analysis methods have not been as successful as expected. Content extraction requires segmentation. Once the segmentation is completed, quantitative data and textures within the boundaries of an object can be easily measured. There is a great need for productive research in texture analysis of real-life patterns.

Image Registration

Image registration in this context is defined as matching images of the same object obtained at different times and under various conditions. Image registration of rigid objects is not difficult to perform. Image registration of non-rigid objects, for example the human body, is difficult to achieve. An example is image registration of a brain anatomical image (MRI) with a physiological image (PET). In this case, the registration is to match anatomical structures of the brain with the physiological activities of each brain structure. Since the brain is encapsulated by the head, a rigid body, the registration task is manageable. On the other hand, the chest and the abdomen do not have a rigid boundary, and involuntary physiological movements can change the location, size, and shape of internal organs. Image registration in this case becomes a very challenging task.

Database Integration and Knowledge Extraction

As more and more data are being generated in health care by various SP procedures, they are stored in various types of databases. Retrieval and knowledge extraction from these databases is difficult. There is an urgent need to develop SP tools for numerous tasks including storage, display, communication, image processing, data compression to support multi-media databases as described in Section 5.10.

This is especially true in two-dimensional image processing. In the health care delivery field, each radiological examination normally generates between 10-40 MB per file. Most of the time only a very small subset of the 40 Mbytes is required for retrieval. A systematic database architecture and novel method for knowledge extraction are needed to accomplish this goal.

6.5 Medical Databases

The National Library of Medicine (NLM) and several thousand medical libraries, covering a broad spectrum ranging from major academic health science center libraries to small rural hospital libraries, constitute a large community entrusted with providing timely and accurate information to practitioners and researchers in biomedicine. Such information can take the form of bibliographic and factual databases, digital images of documents and medical x-rays, video sequences of medical procedures, digitized color images of dermatologic conditions, procedural rules found in expert systems, output of medical devices such as ultrasound, EKG, EMG and many others. These information types reside in many types of storage systems and some of them, mainly bibliographic databases, are routinely accessed via telephone modem and value-added public networks (Telenet, Tymnet). The potential and problems faced by medical databases are indicative of many other database applications.

As a consequence of several trends, professionals (end users) seek easy and rapid access to these diverse resources. These trends include: (a) the rapidly increasing global connectivity of the NREN/Internet; (b) the increasing availability of public domain software, both server (e.g., WWW, Gopher) and client (e.g., Mosaic); and (c) increasing power on the desktop, making true multimedia workstations possible.

Projects that address some of these application goals are being conducted at the NLM's R&D center, the Lister Hill National Center for Biomedical Communications, as at many other universities, major libraries, medical centers and other institutions. The projects at NLM deal with document imaging, x-ray imaging, 3D visualization of anatomic data and others. Document imaging projects include SAIL and DocView. SAIL (System for Automated Interlibrary Loan) is a multi-workstation system connected via a token ring LAN. It accesses the NLM's mainframe computer for interlibrary loan (ILL) requests, parses the requests into document information and user information, passes the document descriptor to a store of about 200,000 already scanned journal page images on optical media, extracts the requested images, and uses an internal fax board along with the user's fax number (previously extracted from the ILL request) to fax the images out to the user's fax machine. The SAIL pilot system is designed to reduce or eliminate the manual activities that are routinely required to fill interlibrary loan requests.

In contrast to SAIL which relies on indirect access to images through the nationwide medical interlibrary loan service, DocView is a direct access system, anticipating the eventual availability of document image databases. It consists of client software running under MS Windows 3.1 that accesses images stored in servers, both inhouse-developed as well as public domain (such as WWW, Gopher, FTP), and retrieves the images over the Internet. DocView enables an end user to preview, display, manipulate, electronically bookmark, segment, cut and paste, and print the document received. DocView has the full capability of communications, document display and manipulation, but it can also use the WWW-Mosaic communications protocol to transfer the images over the Internet, and serve as the necessary viewer or browser.

Since most documents requested by users of their libraries will be in paper form for the foreseeable future, the NLM is also designing a multimodal transmission system that limits the manual activity of an operator to simply scanning a document. This system, called WILL, automatically retrieves ILL requests from NLM's mainframe computer, parses the requests, recognizes the recipient's fax number, mailing address or Internet IP address, and automatically sends the scanned images to the right place by the method requested by the user. The system is designed to place no burden on the operator beyond the scanning.

An x-ray imaging project, DXPNET, is motivated by the 17,000 x-rays taken as part of the nationwide NHANES II survey and 10,000 x-rays taken in the NHANES III survey (NHANES: National Health and Nutrition Examinations Survey). These x-ray films (cervical and lumbar spine; hands and knees) need to be preserved, and furthermore need to be read by radiologists. The NLM is scanning and archiving these in an optical disk jukebox, and building the server and client workstations to allow access to the images over the Internet. The user (radiologist) will be able to access and retrieve the images, and enter numeric readings and free text comments in a form on the screen. These readings and comments will then go to the jukebox where they will be stored in a database for access by other interested parties, e.g., epidemiologists, demographers, insurance companies, and government agencies.

Motion video is also of interest to the biomedical user. The NLM and other medical libraries have repositories of videotapes of surgical procedures, grand rounds, ultrasound examinations, educational seminars and others. The current procedure to gain access to such collections is for a user first to search index databases that provide information about the tapes, such as title, brief textual abstract, general topic, creator, institution, etc. The artifact itself, when requested, has to be packaged and mailed out to the requester. A project is under way in which the tapes are being digitized and compressed via MPEG, stored on optical disks, and accessed automatically by a link from a database search. Potentially important SP methods for video include compression by MPEG or motion JPEG and methods to link up of a citation received from a bibliographic database to a video sequence or some part of a sequence.

SP and Other Tools

Among the diverse technologies required to make medical information resources more accessible are those having to do with efficient storage, improved image quality, application-specific man-machine interfaces, pattern recognition for feature extraction, and efficient transmission.

In the research agenda for document imaging, the following areas are of interest: compression of two-tone images; OCR and image enhancement preprocessing to make it more reliable, such as automated border removal, skew detection and page orientation detection; user interfaces allowing easy manipulation of the images; efficient page description and markup languages; scaling and scale-to-gray algorithms for image quality improvement.

In the document imaging projects, a key element is the scanning that is needed to create a suitable image store. In the process of scanning, SP stages are: automated border (page edge effects) detection and removal, automated page orientation detection, automated skew detection. All of these not only aid the capture of the document images with high visual quality level, but also provide the quality needed for subsequent processing such as accurate optical character recognition. Also, in the display of the document images, other useful processing includes: subsampling scaling (from the scanned 300 dpi image to VGA or SVGA resolution), and scale-to-gray to increase readability by shading the image in suitable gray levels rather than black or white. Ordinary scaling replaces groups of pixels by a single pixel, either black or white, while scale-to-gray replaces them with pixels of intermediate gray values. Ordinary scaling allows fast compression via runlength coding but the characters often appear broken, and readability could be impaired. Scale-to-gray increases readability by shading the character edges, but image compression is slow.

In the case of digitized x-rays and color originals, the following areas are of particular interest: compression of multiple gray level images; design and optimum selection of quantization tables for the DCT process within JPEG; automated and user-directed segmentation based on pattern recognition for feature extraction by mathematical morphology techniques; noise removal by morphologic operators; contrast enhancement by histogram equalization; image classification using statistical classification algorithms. Any algorithm that claims enhancement or useful segmentation must carefully be validated to demonstrate that the enhancement or segmentation are indeed useful to the user radiologist or other MD.

Of particular concern to the ongoing development of the NII is that transmission over the Internet of digitized images is relatively slow, especially for large files such as those resulting from the scanning of x-rays. For example, cervical and lumbar x-rays when scanned result in files approximately 5 and 10 MB respectively. This motivates both improved means of lossless compression, for speeding transmission of exact representations, and improved means for validating lossy compressed images in specific applications. Published research suggests that lossy compression, with reproduction visually indistinguishable from original images, can yield a ratios conservatively estimated at 10:1, five times that of typical lossless compression algorithms. This is clearly applicable in some areas such as recall, education, and research, but it is a matter of controversy in critical applications such as clinical diagnosis. Hence it is important both to improve the quality of lossy compression algorithms and to develop, with the help of engineers, radiologists, and biostatisticians, careful validation procedures to quantify the diagnostic quality and utility of an image.

One possibility for speeding lossless compression is the use of multisocket transmission methods based on the observation that during a conventional single socket linkup the transaction overhead for TCP/IP limits the effective rate at which data can be transferred. However, if the image to be transmitted were to be sectioned, each section assigned a pair of sockets (at the transmission and receiving ends), and the sections transmitted and reconstructed at the other end, the effective transfer rate could be higher. Preliminary tests have shown a three-fold increase in transmission speed for about 20 socket pairs.

Another element in access is the increasingly distributed nature of medical information at many points on the Internet. There is also a trend toward multimedia workstations that enable a user to view, manipulate and seamlessly move among different data types on the desktop. Both these trends motivate intelligent gateways to the appropriate information sources, the extraction of the query-driven data, and reconstruction of the data at the client workstation for presentation to the user for viewing and to possibly do further processing and analysis.

6.6 Education

The accomplishments and potential for SP in education are exemplified by the NSF-sponsored Image Processing for Teaching (IPT) project developed at the University of Arizona, which has pioneered the use of digital image processing as a science and mathematics teaching tool for all grades, from upper elementary through college. Using state-of-the-art software, students and their teachers discover and explore in the same manner as scientists. In addition, the use of images avoids the language- and code-based way science has been traditionally introduced to students. Consequently, more students can be attracted to science and have an opportunity for meaningful participation and learning through exploration and discovery.

The IPT project provides unique constraints on the structure of the NII because it is evidently developing ways for more Americans to use more data in a meaningful and productive way than any other application. Specifically, the targeted user population is 50 million American school students and their teachers; the data include all scientific imaging data, comprising all two-dimensional digital arrays of data from the broad suite of instruments and experiments of recent and future research in all fields. Because most classrooms do not yet have viable access to the Internet, IPT materials and images have so far been distributed via CDROM. As some IPT teachers have found ways to access Internet, they have been able to add an extra dimension of exploration to their students' IPT activities.

For example, several IPT curriculum units involve analysis of GEOS weather satellite imagery of hurricane systems and diurnal surface temperature variations. IPT CDROMs contain time-sequence data sets for these activities as well as supplemental image data for further exploration. Consider the extension of this activity that became possible when an IPT middle-school teacher on the Tohono O'odham reservation gained Internet access: His students were able to obtain same-hour satellite data for correlation with local weather conditions. By using Internet, the IPT activity was extended to a term-long project in weather prediction and satellite data interpretation.

It is critical to recognize that simply making data accessible on the Internet is not enough to ensure that they will be used in education, even as classrooms begin to be connected. These data will only be broadly used by the K-12 educational community if the following support is provided:

1. widespread teacher education,
2. curriculum-based materials development, and
3. substantial on-line follow-up support.

IPT has already developed the infrastructure for broad dissemination of those essential components. Over 1000 teachers have taken at least five-day-long workshops in IPT at sites across the country; curriculum materials are available in CDROM form and allow for expansion by those students who have access to imaging data on the Internet; extensive follow-up support and out-reach are readily available through telephone and e-mail hotlines, conferences, newsletters, and school visits. Thus the IPT project serves as a full-scale model of successful implementation of massive widespread usage of the kinds of information that should be made available on the NII.

Another project exploring the use of image processing in education is entitled "Mathematics Experiences Through Image Processing" which focuses on the matter of motivating middle-school and high-school students to take mathematics seriously. Because typical American children are brought up today on television and video games, they are at home with images and easily attracted to activities involving special effects and visual transformations. A digital image, however, is not only a visual object but also a mathematical one, and the many mathematical operations on images have corresponding visual effects that are quite striking. As with the IPT project, the NII allows these learning experiences to benefit from fresh, authentic imagery from around the world, and it helps build connections between mathematics and ideas from other subjects, as called for in the National Council of Teachers of Mathematics' *Curriculum and Evaluation Standards for School Mathematics*. In the future, the NII can make servers available that can compute image transformations designed by student with high-resolution imagery, enhancing the mathematics/image-processing experiences that students get from their local PCs alone. In this way, the NII itself can contribute to the kind of mathematics education that future US citizens will need in order to be intelligent NII users.

6.7 Employment Services

A challenging aspect of greater automation in manufacturing and services is the question of where new jobs will come from to keep all able minds gainfully employed. Traditional manufacturing firms and even computer makers have been laying off people to reduce costs and, in conjunction with automation, to raise per-worker productivity.

Business spokesmen claim that the bulk of new jobs will grow out of small businesses. The NII can contribute to this growth in at least two ways. First, it can help those seeking jobs to learn of and reach those offering jobs. But probably better than this, it can facilitate the creation of new companies, and subsequently, the creation of new jobs.

The NII can help job seekers in finding jobs. Information services that list jobs in an appropriately searchable way could greatly ease the often long and frustrating search process that many job seekers go through. It is even possible that sophisticated software agents could perform some of the introductory informational transactions between job applicant and employer.

More exciting from a SP point of view is the prospect of using video-conferencing systems to support the kinds of interpersonal contact that lead to team building and formation of small businesses, only at a distance. Small companies often target niche markets where specialization and close matches of employee knowledge to the company's business are key to success. The chances of finding the right people for a viable small business are enhanced if the community of participants is not limited to a local geographical area. However, working with strangers at a distance requires a degree of trust that must be established through relatively intimate kinds of communication. Email and telephone communication may be sufficient in some contexts to establish such a level of trust, but high-quality video conferencing can inspire trust more effectively.

Appendix A

Participants: Addresses

Dr. Julia Abrahams, ONR
Office: (703) 696-4320
fax: (703) 696-2611
email: julia@tomcat.onr.navy.mil
Mathematical, Computer and Information Sciences Division, Code 311
Office of Naval Research
Arlington, VA 22217-5660

Professor George Cybenko, Dartmouth College.
Office: (603) 646-3843
fax: (603) 646-3856
email: George.Cybenko@Darmouth.edu
Thayer School of Engineering
Dartmouth College
8000 Cummings Hall
Hanover, NH 03755-8000

Professor Martin Cohn, Brandeis Univ.
Office: (617) 736-2705
fax: (617) 736-2741
email: marty@cs.brandeis.edu
Computer Science Department
Brandeis University
Waltham, MA 02254

Professor Tom Fischer, Washington State University.
Office: (509) 335-4960
fax: (509) 335-3818
email: fischer@eecs.wsu.edu
Dept. of Electrical & Computer Engineering
Washington State University
Pullman, WA 99164-2752

Professor Allen Gersho, UC Santa Barbara.
Office: (805) 893-2037
fax: (805) 893-3262
email: gersho@ece.ucsb.edu
Center for Information Processing Research
Dept. of Electrical & Computer Engineering
University of California
Santa Barbara, CA 93106

Professor Michael Goodchild, UC Santa Barbara.
Office: (805) 893-8049
fax: (805) 893-7095
email: good@ncgia.ucsb.edu
Director, National Center for Geographic Information and Analysis
Professor of Geography
University of California
Santa Barbara, CA 93106-4060

Professor Robert M. Gray, Stanford University.
Office: (415) 723-4001
fax: (415) 723-8473
email: gray@isl.stanford.edu
Information Systems Laboratories
Stanford University
Stanford, CA 94305

Professor Richard Greenberg, University of Arizona.
Office: (602) 621-6940
fax: (602) 621-4933
email: greenber@hiips.lpl.arizona.edu
Lunar and Planetary Lab,
University of Arizona
Tucson AZ 85721

Dr. Ali Habibi, The Aerospace Corporation

Office: (310) 336-8751
fax: (310) 336-1474
email: habibi@jenner.aero.org
M4-902
The Aerospace Corp.
2350 El Segundo Blvd
El Segundo, Ca 90245

Dr. Barry Haskell, Bell Telephone Labs.
Office: (908) 949-5459
fax: (908) 949-3697
email: bgh@research.att.com
Room HO 4C538
AT&T Bell Labs
Holmdel, NJ 07733-3030

Professor H.K. (Bernie) Huang, UC San Francisco.
Office: (415) 476-6044
fax: (415) 502-3217
email: bernie_huang@radmac1.ucsf.edu
Dept of Radiology, UCSF
505 Parnassus Ave.
San Francisco, CA 94143-0628

Professor John Limb, Georgia Tech.
Office: (404) 894-9106
fax: (404) 894-0272
email: limb@cc.gatech.edu

Dr. Tom Lookabaugh, Divicom.
Office: (408) 944-6704
fax: (408) 944-6705
email: tdl@divi.com
DiviCom, Inc.
580 Cottonwood Drive
Milpitas, CA 95053

Dr. Majid Rabbani, Kodak.
Office: (716) 477-3722
fax:(716)722-0160
email: rabbani@kodak.com
Eastman Kodak Company
1700 Dewey Avenue
Rochester, NY 14650-1816

Dr. Nora Sabelli, NSF
Office: (703) 306-1651
fax: (703) 306-0434
email: nsabelli@nsf.gov
Program Director, Program for Applications of Advanced Technologies
Education and Human Resources, and
Program for Networking Infrastructure for Education
Computer and Information Science and Engineering and
Education and Human Resources
National Science Foundation
Stafford Building Suite 855
4201 Wilson Blvd.
Arlington, VA 22230

Dr. Jon Sjogren, AFOSR
Office: (202) 767-4940
fax: (202) 404-7496
email: jas@eng.umd.edu or sjogren@afosr.af.mil
AFOSR/NM
110 Duncan Ave. Suite B115
Bolling AFB, DC 20332-0001

Professor Tom Stockham, Stockham Technologies.
Office: (801) 583-7777
fax: (801) 583-2608
email: stockham@castor.elen.utah.edu
1100 Vista View Drive
Salt Lake City, Utah 84108-2525

Professor Steve Tanimoto, University of Washington.
Office: (206)-543-4848
fax: (206)-543-2969
email: tanimoto@cs.washington.edu
Dept of Computer Science and Engineering, FR-35
University of Washington
Seattle, WA 98195

Dr. George Thoma, National Library of Medicine.
Office: (301) 496-4496
fax: (301) 402-0341
email: thoma@lhc.nlm.nih.gov

Professor Martin Vetterli, UC Berkeley (and Columbia)
Office: (510) 642-9719
fax: (510) 642-2845
email: martin@wavelet.eecs.berkeley.edu
Office: 267 Cory Hall
Dept. of EECS
UC Berkeley

231 Cory Hall 1770
Berkeley, CA 94720-1770

Appendix B

Participants: Biographies

Julia Abrahams received the B.S. degree in engineering and applied science from Yale University, in 1974, the M.S. degree in electrical engineering from Stanford University, in 1975, and the Ph.D. degree in electrical engineering and computer science from Princeton University, in 1979. She served on the faculty in electrical engineering at Rice University and on the faculty in electrical and computer engineering at Carnegie Mellon University and worked as the engineering editor at Springer Verlag, New York. Since 1988 she has been with the Office of Naval Research where she administers basic research programs in probability and statistics. Her current research interests include source coding, data compression, and combinatorial search.

Martin Cohn is a member of the Computer Science Department of Brandeis University, where he teaches computation theory, algorithms, information theory, and cryptography. He has been Program Chair and Proceedings Co-editor for the Data Compression Conferences (DCC) of 1992 through 1995. Dr. Cohn served as an expert witness in data-compression litigation between Stac Electronics and Microsoft Corp., and he has contributed the data-compression section for the Encyclopedia of Computer Science and Technology. He was a co-author of the first Lempel/Ziv data compression patent.

George Cybenko is the Dorothy and Walter Gramm Professor of Engineering at Dartmouth College. He received his Ph.D. in Electrical Engineering and Computer Science from Priceton University in 1978 and his B.A. in Mathematics from the University of Toronto in 1974. Prior to joining Dartmouth, Cybenko was Professor of Electrical and Computer Engineering at the University of Illinois at Urbana-Champaign where he was also Associate Director of the Center for Supercomputer Research and Development. His current research interest is the application of signal processing and machine learning techniques to novel information retrieval and management problems that arise in networked multimedia systems.

Thomas R. Fischer was born in Fairbanks, Alaska, in 1953. He received the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Massachusetts, Amherst, and the Sc.B. degree magna cum laude from Brown University. From June 1975 until August 1976, he was a Staff Engineer at the Charles Stark Draper Laboratory, Cambridge, MA. From 1979 until 1988 he was with the Department of Electrical Engineering, Texas A&M University, first as Assistant Professor, then as Associate Professor. Since January 1989 he has been a Professor in the School of Electrical Engineering and Computer Science at Washington State University. His current research interests include image and speech compression, digital communications, and digital signal processing. From 1989 to 1992 Professor Fischer was Associate Editor for Source Coding for the *IEEE Transactions on Information Theory.* He is a member and past Secretary of the Signal Processing and Communication Electronics Technical Committee of the IEEE Communications Society. In 1987 Professor Fischer received an outstanding teaching award from the College of Engineering at Texas A&M University. He was a co-recipient of the IEEE Signal Processing Society's 1993 Senior Award in the Speech Processing Area.

Allen Gersho is Professor of Electrical and Computer Engineering at the University of California, Santa Barbara (UCSB). He received his B.S. from MIT in 1960 and Ph.D. from Cornell University in 1963. He was at Bell Laboratories from 1963 to 1980. His current research activities are in signal compression methodologies and algorithm development for speech, audio, image and video coding. He is co-author with R.M. Gray of the book, *Vector Quantization and Signal Compression*, Kluwer Academic Publishers, 1992, and co-editor of two books on speech coding. He is a member of various IEEE technical, award, editorial, and conference management committees He received NASA "Tech Brief" awards for technical innovation in 1987, 1988, and 1992. In 1980, he was co-recipient of the Guillemin-Cauer Prize Paper Award from the Circuits and Systems Society. In 1981, he became a Fellow of the IEEE. He received the Donald McClennan Meritorious Service Award from the IEEE Communications Society in 1983, and in 1984 he was awarded an IEEE Centennial Medal. In 1992, he was co- recipient of the 1992 Video Technology Transactions Best Paper Award from the IEEE Circuits and Systems Society.

Michael F. Goodchild received his BA in Physics from the University of Cambridge in 1965, and his PhD from McMaster University in 1969. He taught at the University of Western Ontario for 19 years, and moved to his present position of Professor of Geography, University of California, Santa Barbara in July 1988. He is currently Director of the NSF funded National Center for Geographic Information and Analysis (NCGIA). His research interests focus on the generic issues of geographic information, including accuracy and the modeling of uncertainty, the design of spatial decision support systems, development of methods of spatial analysis, and data structures for global GIS.

Robert M. Gray received the B.S. and M.S. degrees from M.I.T. in 1966 and the Ph.D. degree from U.S.C. in 1969, all in Electrical Engineering. Since 1969 he has been with Stanford University, where he is currently a Professor and Vice Chair of the Department of Electrical Engineering. His research interests are the theory and design of data compression and classification systems. He is the author or coauthor of five books and coeditor of two collections on related topics. He was an Associate Editor of the *IEEE Transactions on Information Theory* from September 1977 through October 1980 and Editor-in-Chief of that journal from October 1980 through September 1983. He was co-chair of the 1993 IEEE International Symposium on Information Theory and has served on the Program Committee of the annual Data Compression Conference since its founding. He currently is a member of the Image and Multidimensional Signal Processing Technical Committee of the IEEE Signal Processing Society. He was corecipient of the 1976 IEEE Information Theory Group Paper Award and the 1983 IEEE ASSP Senior Award. He received an IEEE Centennial medal in 1984 and the IEEE Signal Processing 1993 Society Award in 1994. He is a Fellow of the IEEE and the IMS.

Richard Greenberg is Professor of Planetary Sciences and Professor of Teaching and Teacher Education at the University of Arizona. Greenberg is Director of the University's Science and Mathematics Education Center. He leads an active research program in planetary science. As a Principal Investigator in NASA's Planetary Geology and Geophysics program, he and his research team carry out investigations of the dynamical evolution of the solar system, including studies of asteroids, meteorites, planetary rings, and the formation of the planets. He is a member of the Imaging Team for NASA's Galileo spacecraft, en route to Jupiter. He directed Galileo's imaging of the Earth during its 1992 fly-by. Since 1989 Greenberg has led the University of Arizona's campus-wide initiative in support of pre-college science, mathematics, and technology education. He also founded and directs the Image Processing for Teaching (IPT) project, giving students in classrooms across the nation the power to engage in substantive scientific exploration and discovery using state-of-the-art digital image processing. He chairs the Board of Trustees of the Center for Image Processing in Education.

Ali Habibi was born in Hamadan Iran in 1942. He received the B.S. at West Virginia University in 1965, the M.S. at Cornell in 1967, and the Ph.D. at Purdue University in 1970, all in Electrical Engineering. During 1970–1971 he was at AT&T Bell Labs,Holmdale,N.J. He was an Assistant Professor of Electrical Engineering at the University of Southern California from 1971 to 1974. From 1974 to 1981 he was a Senior Staff Engr.at TRW. Since 1981 he has been a Senior Engr. Specialist at the Aerospace Corp. He is a Senior Member of the IEEE and a member of SPIE, the Communication Theory Committee of the IEEE Communication Society, Eta Kappa Nu, and Tau Beta Pi. He was Editor (1974–1986) of the *IEEE Trans. Comm.* He received the 1972 Best Paper Award from the IEEE Comm. Soc. His research interests are Image Processing; Digital Signal Processing; Communications; Radar Signal Processing; and Optical Communications.

Barry G. Haskell received the B.S., M.S. and Ph.D. degrees in Electrical Engineering from the University of

California, Berkeley. Since then he has been at AT&T Bell Laboratories, Holmdel, New Jersey and is presently Head of the Visual Communications Research Department. He has also taught graduate courses at Rutgers University, City College of New York and Columbia University. His research interests include digital transmission and coding of images, videotelephone, satellite television transmission, medical imaging as well as most other applications of digital image processing. He has published over 35 papers on these subjects and has 20 patents either granted or pending. He is also the coauthor of the book *Digital Pictures : Representation and Compression*. Dr. Haskell is a member of Phi Beta Kappa and Sigma Xi, and is a Fellow of the IEEE. He is a member of the ITU-T SG15 Experts Group on Videotelephony, as well as the ISO WG11 Motion Picture Experts Group (MPEG). He is an Associate Editor of the *IEEE Transactions on Communications* and the *IEEE Transactions on Circuits and Systems* for Video Technology.

H.K.(Bernie) Huang, D.Sc., FRCR(Hon), Professor and Vice Chairman, Director of Lab. for Radiological Informatics, Department of Radiology, UCSF Medical School. Dr. Huang's research interests are in Medical Imaging and Informatics, Picture Archiving and Communication Systems, high-speed image communication, distributed and parallel computing, and image compression.

John O. Limb joined Bell Laboratories, Holmdel, NJ, in 1967 and became Department Manager, Visual Communications Research Department in 1971. He worked for a number of years on the coding of color and monochrome picture signals to reduce channel capacity requirements and has published widely in this area. He also worked and published in the areas of visual perception, office information systems and local/metropolitan area networks. In 1984 he joined Bell Communications Research and in 1986 he was appointed Director of the Networks and Communications Laboratory at Hewlett-Packard Laboratories, Bristol, England. In June 1989, he returned to the U.S. with Hewlett-Packard as Lab Manager, Technology Analysis, Cupertino, CA. In 1992 he returned to HP Labs as Lab Director, Media Technology Lab. In July 1994 Dr. Limb joined Georgia Institute of Technology to accept the Eminent Scholar in Advanced Telecommunications chair in the College of Computing. Dr. Limb is a Fellow of the IEEE and past Editor-in-Chief of the *IEEE Transactions on Communications* and the *IEEE Journal on Selected Areas in Communications*. He was co-recipient of the 1990 IEEE Alexander Graham Bell Medal.

Majid Rabbani received his Ph.D. degree in electrical engineering from the University of Wisconsin at Madison in 1983. He joined the Eastman Kodak Research Laboratories, Rochester, New York in the same year, where he currently is a Research Associate and the head of the image compression Technology Area. He is also an adjunct faculty at the department of electrical engineering at Rochester Institute of Technology (RIT). Dr. Rabbani is the recipient of the 1988 Kodak C. E. K. Mees Award (Kodak's highest research honor) for excellence in research. His current research interests include multidimensional digital signal processing and image processing where he has published over 50 technical articles and holds 9 patents. He is a senior member of IEEE and a fellow of SPIE. He has coauthored the book *Digital Image Compression Techniques*, published by SPIE Press in 1991 and edited the SPIE Milestone Series on *Image Compression and Coding*, published in 1992.

Thomas G. Stockham, Jr. received the Sc.D. degree in 1959, from MIT. In 1959, he was appointed Assistant Professor of Electrical Engineering. From 1966 to 1968, he was a Staff Member of the MIT Lincoln Laboratory, where he developed nonlinear systems for image enhancement and audio companding using the principles of generalized superposition. In July 1968, he the joined Computer Science faculty at the University of Utah, where he worked in the fields of signal processing and computer graphics. During 1973–1974 he served on a panel of experts for Chief Judge John J. Sirica to examine the Watergate Tapes. In 1975 Dr. Stockham founded Soundstream, Inc. Under his direction the company developed digital commercial sound recording and editing. These developments led the establishment of these practices industry wide. The resulting accumulation of more than 100 digitally recorded albums by early 1981 triggered the advent of the compact disc (CD). In 1983 he joined the University of Utah Department of Electrical Engineering. He is also an active consultant in the field of digital signal processing, especially digital image and sound processing technologies. Dr. Stockham is a Fellow of the IEEE and the Audio Engineering Society and served as president of the latter in 1982–1983. He received the the SMPTE Alexander M. Poniatoff Gold Medal for Technical Excellence in 1985, the Audio Engineering Society Gold Medal in 1987, a National Academy of Television Arts and Sciences *Emmy* in 1988 for the development of tapeless audio recording and editing technology, and a National Academy of Recording Arts and Sciences *Grammy* in 1993 for his pioneering role in the development and advancement of digital recording.

Steven L. Tanimoto is Professor of Computer Science and Adjunct Professor of Electrical Engineering at the University of Washington in Seattle. His interests include parallel image processing, artificial intelligence and educational technology. He has held visiting positions at the University of Paris, Linköping University (Sweden), Kobe University (Japan), the National Institutes of Standards and Technology, and Thinking Machines Corporation. Dr. Tanimoto served an an Associate Editor of *IEEE Transactions on Pattern Analysis and Machine Intelligence* from 1983 to 1986, and as Editor-in-Chief from 1986-1990. His text, *The Elements of Artificial Intelligence Using Common Lisp*, was published by W.H.Freeman in 1990. He served as the program chair for the 1994 International Conference on Pattern Recognition subconference on parallel computation, and as co-program chair of the 1994 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. He received the A.B. from Harvard in 1971 and the Ph.D. from Princeton in 1975.

George R. Thoma received a BS from Swarthmore College, and an MS and PhD from the University of Pennsylvania, all in electrical engineering. He is the Chief of the Communications Engineering Branch of the National Library of Medicine. His R&D interests are in communications engineering and image processing, as applied to both document and digitized xray images. His former private sector experience at General Electric and AII Systems was in satellite communications and satellite-aided navigation.

Martin Vetterli received his Engineering degree from ETH in Zurich, his MS from Stanford and his Doctorate from EPF Lausanne. He is on the faculty at Columbia University and UC Berkeley. He works on signal processing for communications, and wavelet theory and applications.